JOURNAL OF APPLIED PHYSICS

JOURNAL OF APPLIED PHYSICS

Vol. 24, No. 8 AUGUST, 1953

In This Issue

		Page
Set of Principles to Interconnect the Solutions of Physical Systems	Gabriel Kron	965
Contact and Rubbing of Flat Surfaces	J. F. Archard	981
Measurement of Elastic Constants at Low Temperatures by Means of Ultrasonic Waves —Data for Silicon and Germanium Single Crystals, and for Fused Silica	H. J. McSkimin	988
Utilization of Degenerate Modes in a Spherical Cavity	Malcolm R. Currie	998
Nature of Solutions of a Rayleigh-Type Forced Vibration Equation with a Large Coefficient of Damping	Paul Brock	1004
Diffusion of Barium in an Oxide-Coated Cathode	Renate S. Bever	1008
Degradation of Polyisobutylenes on Shearing in Solution	A. B. Bestul and H. V. Belcher	1011
Nonlinear Semiconductor Resistors	F. A. Schwertz and J. J. Mazenko	1015
Filters for Detection of Small Radar Signals in Clutter	Harry Urkowitz	1024
On the Interference of Pulse Trains	K. S. Miller and R. J. Schwarz	1032
Beta-Ray Spectrometer Line Shapes with Tilted Sources	H. W. Kruse, G. P. Mellor, and C. M. Fowler	1037
Measurement of Particle Size of Fine Ferromagnetic Powders	A. D. Franklin, R. Campbell, and J. Weinman	1040
Determination of Reflection Coefficients and Insertion Loss of a Wave-Guide Junction	Georges A. Deschamps	1046
Erratum: Determination of the Drag on a Cylinder at Low Reynolds Numbers	R. K. Finn	1050
Replica Technique in the Study of Chemical Precipitation Processes	Robert B. Fischer and Joseph P. Ellinger	1051
Methods of Measuring the Properties of Ionized Gases at High Frequencies. IV. A Null Method of Measuring the Discharge Admittance	Lawrence Gould and Sanborn C. Brown	1053
Electrical Properties of Liquid Selenium I	H. W. Henkels and J. Maczuk	1056
Letters to the Editor:		
Schlieren Photographs of Sound Fields	Daniel S. Schwartz and Anthony L. Russo	1061
Effect of Cold Work and Anneal on Resistivity of Alpha Brass	Herbert I. Fusfeld	1062
Simpler Explanation for the Observed Shot Effect in Germanium Filaments	A. van der Ziel	1063
Theory of the Flicker Effect	T. B. Tomlinson and W. L. Price	1063
Noise in CW Magnetrons	David Middleton, W. M. Gottschalk, and J. B. Wiesner	1065
Effect of Halogenation on Frictional Properties of Plastics	R. C. Bowers, W. C. Clinton, and W. A. Zisman	1066
Comment on the Temperature Dependence of the Viscosity of Liquids	Arthur K. Doolittle	1067
Further Note on the Temperature Dependence of the Viscosity of Liquids	F. Gutmann and L. M. Simmons	1067
Summation of Slowly Converging Series	Igor Gumowski	1068
Index to Advertisers		xxiv

ELMER HUTCHISSON, *Editor*

EARLE C. GREGG, *Assistant Editor*

Board of Editors
RAY E. BOLZ
LYLE D. BORST
WALLACE J. ECKERT
ROBERT D. HEIDENREICH
LEONARD B. LOEB
W. JAMES LYONS
LOUIS A. PIPES
ROMAN SMOLUCHOWSKI
W. R. SMYTHE

Publication of the American Institute of Physics

OFFICERS
GEORGE R. HARRISON, *Chairman*
GEORGE B. PEGRAM, *Treasurer*

GOVERNING BOARD
J. W. BUCHTA
KARL K. DARROW
G. J. DIENES
S. A. GOUDSMIT
GEORGE R. HARRISON
FREDERICK V. HUNT
DEANE B. JUDD
HUGH S. KNOWLES
C. C. LAURITSEN
WILLIAM F. MEGGERS
PHILIP M. MORSE
BRIAN O'BRIEN
GEORGE B. PEGRAM
I. I. RABI
DUANE ROLLER
FREDERICK SEITZ
WILLIAM SHOCKLEY
RICHARD M. SUTTON
J. H. VAN VLECK
MARK W. ZEMANSKY

Consultant on Printing
MELVIN LOOS

ADMINISTRATIVE STAFF
HENRY A. BARTON, *Director*
WALLACE WATERFALL, *Executive Secretary*
THEODORE VORBURGER, *Advertising Manager*
RUTH F. BRYANS, *Publication Manager*
BETTY H. GOODFRIEND, *Assistant Secretary*
EILEEN B. NEUBERGER, *Circulation Manager*
KATHRYN SETZE *Assistant Treasurer*

The JOURNAL OF APPLIED PHYSICS, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as PHYSICS. The JOURNAL OF APPLIED PHYSICS publishes editorials and reviews, as well as technical papers of applied physics. Articles appearing in it are indexed regularly, according to their field, by Physics Abstracts, Chemical Abstracts, Engineering Index, Inc., Industrial Arts Index, etc.

Manuscripts should be submitted to Elmer Hutchisson, Editor, Case Institute of Technology, Cleveland 6, Ohio. Unless otherwise stated, submission of a manuscript is a representation that it has been neither copyrighted, published, nor submitted for publication elsewhere. The authors' institutions are requested to pay a **publication charge** of $8 per page which, if honored, entitles them to 100 reprints without further charge. Instructions will be sent with galley proofs.

Subscriptions, renewals, and orders for back numbers should be addressed to the American Institute of Physics, 57 East 55 Street, New York 22, New York.

Subscription Price	U.S. and Canada	Elsewhere
To members of the American Institute of Physics and Affiliated Societies	$10.00	$11.00
To all others	12.00	13.00

Back Number Prices

Yearly back number rate when complete year is available: $14.00. Single copies: $1.25 each.

Changes of address should be sent to the American Institute of Physics; **corrected proofs** should be sent to the Publication Manager, American Institute of Physics. **Advertising rates** supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879. Accepted for mailing at the special rate of postage provided for in paragraph (d-2), section 34.40, P. L. & R. of 1948, authorized May 2, 1932.

Journal
of
Applied Physics

Volume 24, Number 8 August, 1953

A Set of Principles to Interconnect the Solutions of Physical Systems

GABRIEL KRON*
General Electric Company, Schenectady, New York
(Received December 31, 1952)

A set of principles and a systematic procedure are presented to establish the exact solutions of very large and complicated physical systems, without solving a large number of simultaneous equations and without finding the inverse of large matrices. The procedure consists of tearing the system apart into several smaller component systems. After establishing and solving the equations of the component systems, the component solutions themselves are interconnected to obtain outright, by a set of transformations, the exact solution of the original system. The only work remaining is the elimination or solution of the comparatively few superfluous constraints appearing at the points of interconnection.

The component and resultant solutions may be either exact or approximate and may represent either linear or, with certain precautions, nonlinear physical systems. The component solutions may be expressed in numerical form or in terms of matrices having as their elements real or complex numbers, functions of time, or differential or other operators, etc. Boundary value, characteristic value, and time-varying problems of partial differential equations, as well as problems in ordinary differential equations and algebraic equations may be solved in this manner.

The method shown may be extended and generalized so that one can tear apart and afterward reconstruct the solution of extremely large or highly intricate physical systems, often without calculating any inverse matrices at all and always without carrying along unduly large matrices. This extension and generalization of the method is analogous to building skyscrapers by erecting first a steel framework and only afterward filling the gaps between the girders as needed. Those versed in the science of tensorial analysis of interrelated physical systems on the one hand and of large electrical networks on the other, should thereby be able to solve, with the aid of already available digital computers, highly complex physical systems possessing tens of thousands and, in special cases, even hundreds of thousands of variables. The accuracy of machine calculations, of coding and even the correctness of the analytical procedure itself may be simultaneously checked by physical tests at various stages of the computation. The saving in computing time is considerable even in smaller problems; by tearing a physical system into n parts, the usual machine calculations are reduced, in matrix inversion for instance, to a fraction of about $2/n^2$.

The present paper develops in detail the solution of a simple boundary value problem of Poisson's equation. A numerical example of interconnecting the solutions of large electric-power transmission systems appears in reference 3. Many simpler numerical examples are worked out in reference 1.

KEY PRINCIPLES FOR INTERCONNECTING SOLUTIONS

Complete Set of Forces and Constraints

SEVERAL keys had to be fitted into the right keyholes before the riddle of interconnecting solutions of physical systems in general, and of systems with very large number of variables in particular, could be unlocked.

One key is the realization that the mere availability of several sets of equations does not automatically allow one to interconnect their solutions. Besides the generalized forces and their equations that are usually available, it is necessary to have access to all the constraints, all the complementary equations of constraints, and moreover to all the mutual relations existing between the forces and constraints.

Expressed in another way, the defining "equations of state" of any physical system call forth the existence of as many equations as there are total number of forces and constraints in the system. Similarly the "equations of solutions" imply the existence of the same number of equations. The elimination of some or all of the superfluous constraints and the consequent reduction in the number of equations (of state, or of solution) represents already a deviation from a complete mathematical model of the physical system, hence an obstacle to a complete freedom of manipulation.

Electric Circuit Models

A second key is the discovery that the electric circuit models of the equations of mathematical physics (be they algebraic, differential, or integral equations) do establish automatically these missing variables, missing equations, and missing mutual terms between the two

* Consulting Engineer, General Electric Company.

types of variables, at the same time when the force equations alone are modeled. However this automatic generation of complementary equations is maintained only if the process of modeling follows certain definite tensorial principles.

Of course the electric circuits exist only on paper and need not be physically realizable; they themselves are only easily visualizable models of more basic topological concepts underlying the whole method of reasoning.

"Orthogonal" Networks

A third key is the discovery of the "orthogonal" nature of all electric (and other types of) networks, possessing always nonsingular (square) transformation matrices that can transform all the forces and constraints among themselves before solution, as well as after solution. By utilizing not only the configuration space of the generalized force variables, but also the configuration space belonging to the constraint variables ("orthogonal" to the space of force variables) the correct amount of elbow room is provided to pass freely and easily from the complete "equations of state" to the complete "equations of solution" and vice versa at any stage of the development.

Tearing Apart and Interconnecting Physical Structures

The most important complementary keys of the riddle are the concept of "tearing apart" of a complex system into an arbitrary number of component systems and the inverse concept of their "interconnection" into the original system, or into any other possible system.

In this connection it should be mentioned that the science of topology, which deals with the properties of interconnected spaces, had its origin in the researches of Kirchhoff concerning the passage of electric currents through a network. It would therefore appear as if the present study would only be a practical application of some branch of topology. That, however, is not obvious. Topology actually investigates only such properties of networks that remain invariant under bending, stretching, or distortion ("rubber geometry"). Properties that remain invariant under tearing a network apart into isolated portions are outside the sphere of interest of topology (having no metric).

In contrast to the basic interest of topology, the whole theory of tensorial analysis and solution of complex physical systems—as developed by the author during the last quarter century—is based chiefly upon the study of concepts that remain invariant when a network of spaces is broken up into its ultimate component spaces and afterward reconnected into all possible configurations, including that of the original system that was torn apart. This radically new point of view formulates that it is possible to pass freely and easily from the equations of any one possible configuration to the equations of any other possible configuration of the same component physical systems (or component spaces) by means of a set of nonsingular transformation matrices that form a "group," the so-called "connection" group.

The representation of this group is what the author calls in his writings the "connection tensor" $C_{\alpha}^{\alpha'}$, containing usually ± 1 or 0 as its components for all stationary networks. The author is not aware of the existence of this concept in topology, although the latter swarms with other types of matrices having ± 1 and 0 as their components, such as the "incidence" matrix, etc. If a metric is introduced into topology, then the concept of invariance under tearing may become a proper subject of investigation. However, the author is not aware of the existence of such a geometrical science of tearing.

The expression "equation" includes either the defining equations of performance (equations of state) or the equations of solutions, both sets containing all the forces and constraints of the system, without any of them being eliminated. Of course practical requirements do necessitate the elimination of superfluous variables. However, the latter may be accomplished in such an organized manner that it does not handicap the above-mentioned freedom of action and transformation at any step of the development.

Tensor Analysis

Still another key is the use of tensor analysis in circuit studies, that is the use of the concepts of "transformation," "invariance," and "group" in the analysis of complex stationary structures. Tensor analysis has been applied by other writers on classical physics only to field problems and to dynamics, but not to circuit analysis.

The concepts of tensor analysis are indispensable in interconnecting solutions of actually existing physical systems. The reference frames needed in the interconnected systems are far different from those in the component systems, and the series of transformations necessary to pass from one frame to another are often quite involved and not at all obvious. Also the reference frames that occur when the defining "equations of state" are interconnected are quite different from those appearing when "equations of solutions" are connected together.

The immense variety of informations that are expected to be deduced from the "solutions" (in conformity with the various restrictions of the problem) offer unlimited opportunities to impinge upon the method of attack the potentialities inherent in the modern mathematical concept of "invariance." Using the terminology of the electric circuit models, sometimes the total "power input" into the system remains invariant or only part of the input. Often a group of impressed voltages only have to remain invariant, in some cases the input impedances must, and so on. Depending upon the type of problem and the nature of the boundary

conditions, examples of invariance could be continued indefinitely.

The basic importance of tensor analysis in the present method may be appreciated from the single fact that it is impossible to establish an electric circuit model (or any other physical model) for a set of equations describing a physical phenomenon, unless the descriptive equations are "tensor" equations (or "tensor-density" equations).

Matrix Routine in Manipulations

The last key to be mentioned is the representation of the components of tensors along any particular reference frame in the form of n-way matrices, in particular in terms of only one-way and two-way matrices. Digital computer routines are best expressed in terms of matrices; also all manipulations on paper can be performed the fastest if "boxes" with a ruled grid are drawn.

This matrix routine for tensors has been employed by the author from the very earliest of his researches, even though practically no other author on tensor analysis uses it. Writers on tensors rarely work out any problem and then only express the components of a tensor along any frame as a "set" of expressions, differentiated from each other only by their proper indices. The matrix routine adapts the method of tensors from a tool for generalized thinking to a tool for getting also numerical answers quickly and systematically.

Summary

When the above-mentioned pieces of the jigsaw puzzle are put together, the following unified view of a particular physical structure appears:

(1) It is possible to pass freely from any isolated portion of the system to any other isolated portion, or to the whole system itself in any direction. (2) It is possible to pass freely from one interconnected structure to any other interconnected structure containing the same component parts in any combination. (3) It is possible to pass freely from the "equation of state" of any structure or its parts to its "equation of solution" in any direction and combination. (4) It is possible to pass freely from any type of boundary conditions to any other.

That is all barriers between "equations of state" and "equations of solution" of a group of physical systems interconnected in arbitrary manner have disappeared and a highway between these hitherto isolated physical concepts has been constructed. The highway is wide enough for utilizing all the capabilities of modern electronic digital computers already constructed or still contemplated.

The physical systems that are to be interconnected may differ radically from each other. They may cover from electrical and mechanical through thermal and chemical phenomena the whole gamut of physics in any combination.

Although the present paper emphasizes the solution of physical systems only, it should be stated that all nonphysical problems which can be represented by tensor equations may also be torn apart, solved in pieces, then reconnected. Extensive problems of the calculus of variations for instance (including economic problems and those of "operations research") are ideally suited for solution in easy stages by the method of tearing. However for the piecemeal solution of nonphysical problems an extension of the concepts of this paper is necessary.

The Indispensable Key Concepts

It should be emphasized that in problems with a comparatively small number of variables not all of the key concepts stated are necessary to interconnect solutions. In particular only the following key concepts are indispensable:

(1) The concept of *tearing apart* systems into a set of independent unrelated systems as a "group" of transformations.

(2) The concept of interconnecting smaller systems into larger systems as the "inverse" elements of the same group of transformations.

(3) The concept of availability of all constraints that appear at the points of tearing; also the availability of the complementary equations of these constraints with all their mutual relations with the forces retained.

(4) The concept of orthogonality of forces and constraints, that is the recognition of their dual behavior during the transformations. This recognition is important of course in other problems besides the interconnection of solutions.

The establishment of an electric circuit (or any other) model is only an auxiliary aid and not an absolute necessity. Every physical system may serve as its own model. Those acquainted with a physical system should be able to determine the right constraints and establish the correct complementary equations without any auxiliary model. The electric circuit model serves only as a universal language, as a common denominator of a large variety of phenomena and as a convenient scheduling routine for a larger number of variables. In many problems only portions of the model need be in circuit form.

Similarly the use of tensor analysis as such need not be recognized. The correct interpretation of physical phenomena is equivalent to the adoption of at least elementary tensorial methods, whether it is so acknowledged or not. However, as soon as practical problems are attacked that are beyond the power of analysis of conventional textbooks, without a conscious and deliberate application of tensorial methods the analysis sooner or later collapses under its own weight.

As the number of variables increases, the use of matrices also becomes a manipulative tool of increasing value.

Interconnection of Nonlinear Systems

It should be noted that this paper does not claim to be able to solve the nonlinear equations of component systems. This paper only claims that if the solutions of the component systems are available (no matter what method has been used for their solution) and if the original system has been cut in such a manner that a relation may be established between the constraints appearing at both sides of the cut, then the component solutions may be interconnected to give the solution of the resultant system.

Nonlinear systems can usually be cut at places at which the constraints are related either linearly or at least in a solvable nonlinear manner (the inverse of the matrix of transformation is calculable). When for instance, a nonlinear partial differential equation is replaced by an electric circuit model in which the impedances are functions of the unknown currents or voltages, the circuits can always be cut in such a manner that the currents and voltages at both sides of the cut can be equated to each other. Thereby the matrix of transformation can always be made linear. Similarly control systems with nonlinear elements[1–3] offer unlimited opportunities for linear cuts.

It is true that after the interconnection it is still necessary to eliminate the constraints appearing at the cuts, and these constraints will usually be related to each other and to the retained forces in a nonlinear manner.

Mass Production of "Solutions"

The tensorial method of attack implies a never ending succession of labor-saving devices. Many of which are almost self-evident, but most of them assume a familiarity with the physical behavior of the system itself and with the analytical tool employed.

An obvious labor-saving device appears when all subdivisions are identical, or assume only a few repetitive forms. In such a case one component solution is applicable to a large number of subdivisions. Similarly, the process of interconnection is often identical for a large number of analogous groups.

The solution of systems may be stored away either in numerical form or in the form of structural tensors, so that they may again be interconnected later on—as occasions arise—into the solutions of still larger systems. That is the suggested method of attack carries with it all the potentialities of standardization and mass production in solving problems. Standardized solutions may now be stored away and then reused in many different types of problems; a procedure that cannot be followed with other methods of solutions.

Tensorial Solutions as Evolutionary Organisms

The primary concepts of "tearing apart" and "interconnection" imply potentialities of far reaching extensions.

When a complex system has already been calculated and afterward certain portions have been altered, it is possible to perform the analogous alterations on the corresponding portions of the solutions also. ***There is no necessity to start the analysis and the solution of the altered system all over again from scratch.*** Similarly, when an already solved system is augmented, an analogous extension may be performed on the respective portions of the solutions, leaving the rest of the solutions undisturbed. The alterations in the system may involve other properties besides physical size.

In general the suggested method of solution of physical systems can parallel the same type of growth and evolution that the systems to be studied undergoes, provided the solutions have been correctly organized from a tensorial point of view.

Continuous Checks During the Calculations

In tearing an extensive physical system with thousands of variables into a large number of subdivisions, the accompanying analytical and numerical work offers numerous pitfalls for all types of mistakes. It would be too risky to wait until the very end of the lengthy calculations to tell if something went wrong along the way.

As the method of attack of this paper consists of establishing a series of actually existing physical systems, each intermediary physical structure may be tested by Kirchhoff's laws in terms of the original constants used at the start. The tests will show immediately if mistakes have been made by the computing machine or errors in the coding of the problem, or if the analytical procedure itself was at fault.

Reduction in Matrix-Inversion Time

Although one of the purposes of tearing systems apart is to solve on a computer large problems that cannot otherwise be solved at all, the method of tearing offers many advantages even in simpler problems that could as well be solved as one unit. Let the comparatively elementary problem of solving explicitly a set of simultaneous linear equations by matrix inversion be considered. As is well known, the computer time involved increases as the cube of the number of equations. The two heavy lines on Table I represent approximately the inversion time, on the Univac and on the CPC, respectively, of matrices with three-hundred or less rows and columns.

[1] (a) G. Kron, *Application of Tension to the Analysis of Rotating Electrical Machinery* (General Electric Review, Schenectady, 1938). (b) G. Kron, *Tensor Analysis of Networks* (John Wiley and Sons, Inc., New York, 1939). (c) *A Short Course in Tensor Analysis for Electrical Engineers* (John Wiley and Sons, Inc., New York, 1941). (d) G. Kron, *Equivalent Circuits of Electric Machinery* (John Wiley and Sons, Inc., New York, 1951).

[2] G. Kron, (a) Part I. "The Six basic reference frames," Trans. Am. Inst. Elec. Engrs. **70**, 1239–1246 (1951); Part II. "Off-nominal turn ratios, **71**," 505–512 (1952); (c) Part III. "The primitive division," **71**, 814–821 (1952); (d) Part IV. "The interconnection of transmission systems" (to be published).

[3] Glimn, Kirchmayer, and Stagg, Trans. Am. Inst. Elec. Engrs. **71**, 796–803 (1952).

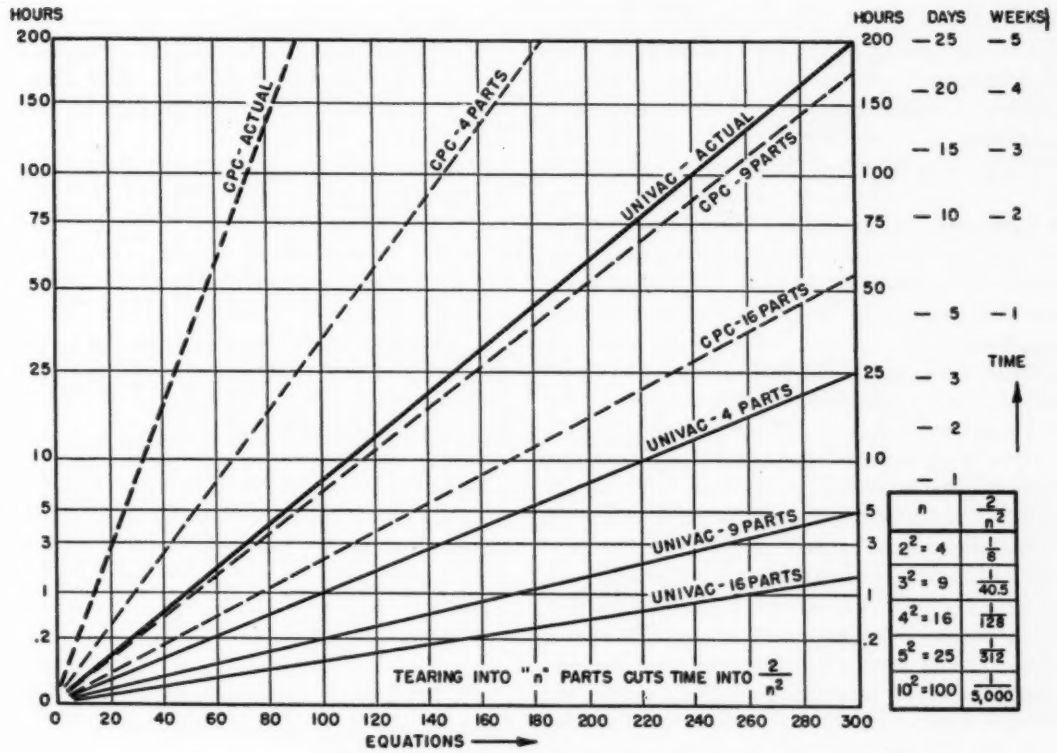


TABLE I. Computer time for matrix inversion.

If a physical system is divided into n parts, the inversion of each part consumes $1/n^3$ of the inversion time of the original system. Hence the n inversions consume $1/n^2$ of the original time. The interconnection consists only of additions, while the solution of constraints may consume about the same time as the inversions. Hence the total solution time is of the order of $2/n^2$ of the conventional inversion time. Moreover, the new type of inverse matrix (the "orthogonal" inverse) contains only a small fraction of nonzero elements (say $1/\sqrt{n}$) existing in a conventional inverse matrix.

For the same accuracy the method of tearing requires also far fewer significant figures to carry throughout the entire calculation.

ELECTRIC CIRCUIT MODELS

Electric Circuits as Topological Models

The equations of a physical phenomenon will serve here to establish an electric circuit model, and it is the model that will be solved. Evidently, from a strictly scientific point of view, the model of a physical phenomenon should be either an algebraic or a geometrical (topological) structure. Since the author is an electrical engineer, and not a mathematician, nor a topologist, he must express his ideas in the medium of which he is most familiar ("the God of dogs barks"). The medium of electric circuit models possesses the virtue of visualizability and practicality, also it may represent physical systems with odd shapes and complications difficult to describe mathematically.

Of course the establishment of an electric circuit model is not an absolute necessity. It is only a stratagem to replace the writing down of a tremendously large number of equations and manipulating them. Other stratagems for accomplishing the same purpose are possible and undoubtedly will be developed in the course of time for various types of problems. Group theory and symbolic logic among others might open up further possibilities in solving problems with perhaps millions of variables, when the electric circuit model will need supplementary supporting frame-works besides tensors.

Electric Circuits as Physical Models

Electric circuit models may represent any partial differential equation of mathematical physics, expressed along arbitrary curvilinear reference frames and along two or three, and even four to five dimensions. Heat flow problems or the field equations of Maxwell, elastic field problems or Schrödinger's wave equation (both the time-varying or stationary-state equations), compressible fluid flow problems in the hodograph plane, or the diffusion of neutrons in an atomic reactor are equally representable with linear circuits. (The author has already published most of the relevant linear electric circuits.[4–9]) All ordinary differential equations can also

[4] G. Kron, Phys. Rev. 64, 126–128 (1943).
[5] G. Kron, Proc. Inst. Radio Engrs. 32, 289–299 (1944).
[6] G. Kron, J. Appl. Mech. 11, 149–161 (1944).
[7] G. Kron, Phys. Rev. 67, 39–43 (1945).
[8] G. Kron, J. Aeronaut. Sci. 12, 221–231 (1945).
[9] G. Kron, J. Appl. Phys. 16, 172–186 (1945).

be represented by electric circuit models, including those of the dynamical equations of Lagrange[1d] and Hamilton (time-varying or stationary-state equations). Linear algebraic equations offer an inexhaustible field for their solution with electric circuit models, such as those of elastic structures composed of large numbers of beams,[10] the representation of valence forces between atoms of polyatomic molecules,[11] electric power transmission systems of several companies operating as a Power Pool,[2] etc. Simultaneous field and circuit problems yield easy representation and solution by means of electrical models.[1d]

The electric circuit models need not be physically realizable. They may contain asymmetrical mutual impedances, negative resistances, phase-shifters,[1d] amplifiers,[12–14] also mechanical networks, electronic and electro-mechanical devices, actual or schematic only. Moreover the branches of the electric circuit models may be two- or higher-dimensional spatial devices, with some of the branches rotating or oscillating with respect to each other by means of moving or sliding contacts. The n-dimensional branches of the model may be Euclidean, Riemannian, or non-Riemannian spaces in any combination and may be expressed along holonomic or nonholonomic reference frames (for instance the branches may contain rotating electrical machines and gyroscopic devices). In many cases the physical system itself may serve as the model (for instance an electric power transmission system[2] or multi-electrode vacuum tube circuits,[12] or groups of rotating electrical machines.[1a])

The present paper is restricted to the solution of a boundary value problem of the simplest type of partial differential equation, that of Poisson's equation. The electric circuit model contains stationary branches with one-dimensional positive resistance elements all arranged in one plane only, without crossing. The reference axes are orthogonal and curvilinear, but otherwise unspecified. The boundary conditions assumed are also the simplest possible, in order to separate the process of "interconnection of solutions" from the process of "introduction of boundary conditions." As the complexity of the model and boundary conditions increases, both processes require further extensions. Nevertheless this simple illustration contains many of the relevant concepts of the new method.

Network and Tensor Theories

Electric circuit models of physical problems serve as ideal vehicles to focus upon the problem at hand certain theories of topology, differential geometry, and tensor analysis that have been applied during the last few decades to the solution of electric power system problems. It will be assumed that the reader is willing to acquire a nodding acquaintance with the tensorial method of analyzing stationary electric circuits, as given in the pages of the book *Tensor Analysis of Networks*.[1b] Constant reference will be made to this book as TAN.

Some aspects of the present theory have already appeared in a series of articles in the *Transactions of the American Institute of Electrical Engineers*, entitled "Tensorial analysis of integrated transmission systems."[2] References will be made to it as TAIT for readers interested in more detailed descriptions and applications of these new ideas. As a matter of fact, the new terminology to be introduced here will follow that used for the solution of integrated electric-power transmission systems. A numerical example is worked out in detail in reference 3.

It may be stated that "all methods given in the pages of TAN and in the other works of the author for interconnecting mesh, junction, and orthogonal networks and equations are equally valid also if the networks (or equations) represent the solutions of physical systems, not only their descriptive equations." Nevertheless networks representing "solutions" require in general far more steps of transformations to interconnect them than the networks that represent merely "equations of state." The additional transformations are the result of the more severe reference-frame requirements of the solutions and of their accompanying boundary conditions.

Orthogonal Networks

Every electric circuit consists of meshes and junction-pairs (or node-pairs). The sum of the number of meshes and junction-pairs is equal to the number of "coils" (branches with nonzero impedances) in the network (TAN, page 73).

In analyzing a network it is customary to write either a set of *mesh* equations $\mathbf{e}=\mathbf{zi}$ or a set of junction-pair equations $\mathbf{I}=\mathbf{YE}$ (TAN, page 354):

$$\begin{bmatrix} \\ \mathbf{e} \\ \\ \end{bmatrix} = \begin{bmatrix} \mathbf{i} & \\ \mathbf{z}_1 & \\ & \end{bmatrix} \Bigg| \Bigg| \begin{bmatrix} & \mathbf{E} \\ \\ \mathbf{I} & \\ \end{bmatrix} = \begin{bmatrix} & \mathbf{E} \\ & \\ & \mathbf{Y}^4 \end{bmatrix}. \tag{1}$$

However, it was shown by the author that for a more general study of any network it is necessary to write as many equations as there are both meshes and junction-pairs (that is, as there are coils). The network is called in that case an "orthogonal" network (TAN, page 407). This extra large number of orthogonal equations may assume either of the following two forms:

$$\begin{bmatrix} \mathbf{e} \\ \mathbf{E} \end{bmatrix} = \begin{bmatrix} \mathbf{i} & \mathbf{I} \\ \mathbf{z}_1 & \mathbf{z}_2 \\ \mathbf{z}_3 & \mathbf{z}_4 \end{bmatrix} \Bigg| \Bigg| \begin{bmatrix} \mathbf{i} \\ \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{e} & \mathbf{E} \\ \mathbf{Y}^1 & \mathbf{Y}^2 \\ \mathbf{Y}^3 & \mathbf{Y}^4 \end{bmatrix}. \tag{2}$$

In the most general study of networks both forces and constraints may occur along any reference axis, so that the equations may assume the following most

[10] G. Kron, J. Franklin Inst. **238**, 400–442 (1944).
[11] G. Kron, J. Chem. Phys. **14**, 19–31 (1946).
[12] G. Kron, Elec. Eng. **55**, 1220–1242 (1936).
[13] G. Kron, J. Appl. Mech. **19**, 107–124 (1948).
[14] G. Kron, Elec. Eng. **67**, 672–684 (1948).

general form:

$$\begin{bmatrix} \mathbf{E}_1+\mathbf{e}_1 \\ \mathbf{E}_2+\mathbf{e}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{z}_1 & \mathbf{z}_2 \\ \mathbf{z}_3 & \mathbf{z}_4 \end{bmatrix} \begin{bmatrix} \mathbf{i}^1+\mathbf{I}^1 & \mathbf{i}^2+\mathbf{I}^2 \end{bmatrix} \quad \Bigg| \quad \begin{bmatrix} \mathbf{i}^1+\mathbf{I}^1 \\ \mathbf{i}^1+\mathbf{I}^2 \end{bmatrix} = \begin{bmatrix} \mathbf{Y}^1 & \mathbf{Y}^2 \\ \mathbf{Y}^3 & \mathbf{Y}^4 \end{bmatrix} \begin{bmatrix} \mathbf{E}_1+\mathbf{e}_1 & \mathbf{E}_2+\mathbf{e}_2 \end{bmatrix}. \tag{3}$$

Equations (1)–(3) may also be written as (TAN, page 414)

$$\mathbf{E}+\mathbf{e}=\mathbf{z}(\mathbf{i}+\mathbf{I}), \qquad \mathbf{i}+\mathbf{I}=\mathbf{Y}(\mathbf{E}+\mathbf{e}),$$
$$E_\alpha+e_\alpha=z_{\alpha\beta}(i^\beta+I^\beta), \quad i^\alpha+I^\alpha=Y^{\alpha\beta}(E_\beta+e_\beta). \tag{4}$$

Each of these two equations may be transformed into each other. However their two sets of component equations (the mesh and junction-pair equations) are independent of each other and in general one set is not derivable from the other set.

Difference Equations and Network Models

Let the partial differential equation of Poisson, $\mathbf{I}=\text{div grad}E$ be replaced by a set of difference equations $\mathbf{I}=\mathbf{YE}$. Their conventional representation, shown in Fig. 1(a), possesses not only junction-pairs but also meshes. The original equations also imply the existence of a set of complementary identities, curl grad$E=0$ as shown in Fig. 1(b), which however does not introduce the missing constraint variables e_α and i^α.

On the other hand, if an electric circuit model is established, as in Fig. 1(c), it is possible to extend the given difference equations along two directions: (1) More columns may be added to the **Y** matrix by introducing mesh impressed voltages $\mathbf{e}=e_\alpha$ (such variables will occur during the process of interconnection). (2) More rows may be added by introducing mesh currents $\mathbf{i}=i^\alpha$ also. These additional rows also contain added columns. (These variables will also appear during the interconnection of solutions.)

Expressed in another way, the original differential equation supplies enough information to establish a model with "junction-pair" properties [Fig. 1(a)] or with "mesh" properties [Fig. 1(b)]. But the differential equation does not suggest the existence of a model with "orthogonal" properties, having four sets of variables **E**, **I**, **e**, and **i**, and having mutual impedances between the mesh variables **e**, **i** and junction variables **E**, **I**. Nevertheless it is this orthogonal property of the electric circuit model, the simultaneous presence of the extra set of mesh variables and of the mutual terms between the original and extra variables—all additional creations of the new model and not an attribute of the original model (the difference equations)—that will enable the introduction of a simplified method of solution.

It may be mentioned that in the potential flow of a nonviscous fluid, where both sources and vortices may appear at singular points, all four variables **E**, **I**, **e**, and **i** may exist in a limited sense (reference 8, Fig. 10). The flow equations however do not supply any mutual terms between the forces and constraints.

Present-Day Methods of Attack

The goal of the analysis is to establish the final **Z** matrix of a very large network whose final performance is defined by the equation $\mathbf{E}=\mathbf{ZI}$. If the values of **I** (impressed sources) are assumed to be known, then the values of **E** (absolute potential of junctions) may be found by merely multiplying a large matrix **Z** by a vector **I**. The components of **I** may assume a large variety of values.

However, the partial differential equation is of the form $\mathbf{I}=\mathbf{YE}$ and not $\mathbf{E}=\mathbf{ZI}$. It does not give the final **Z** matrix, only its inverse, the **Y** matrix. It was always known that the inverse of a very large matrix **Y** could be found by means of submatrices, that is by subdividing the resultant matrix **Y** into $2\times2=4$, or nine, or sixteen, etc., smaller matrices. However this manipulation involves not only the inverse calculation of submatrices, but a very large number of additional matrix multiplications that grows to a prohibitively large number even by a 3×3 subdivision (TAN, page 260, Eq. 10.50).

Although it was also always possible to find the inverse **Z** of each subdivision, it was not practicable to combine the component **Z**'s (instead of **Y**'s) into the resultant **Z** in an exact manner. Various approximate methods have been suggested to piece together the component **Z** solutions either by guessing the final results, or by making appropriate assumptions, or by introducing convolution integrations, and so on.

The New Method of Attack

The new method accomplishes the interconnection of the component **Z** matrices by a routine procedure. The new steps are as follows: (1) Given a physical system with a set of generalized force equations $\mathbf{I}=\mathbf{YE}$. It is

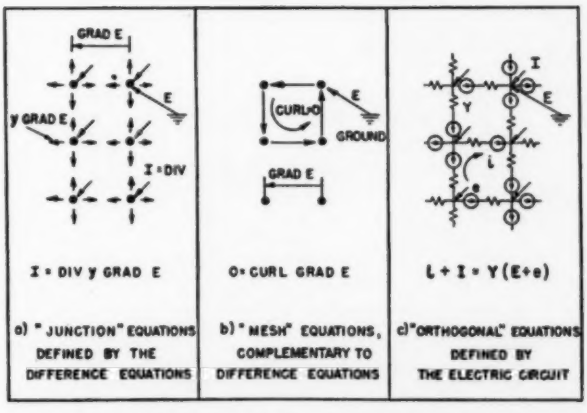


FIG. 1. Representations of Poisson's equation.

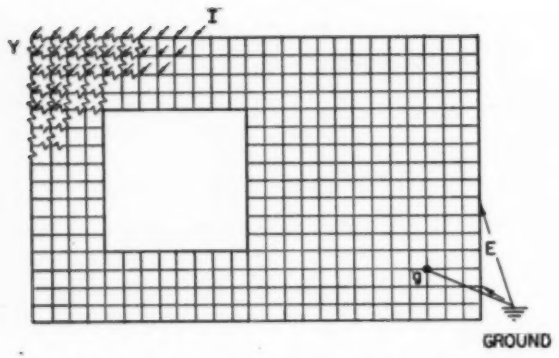


FIG. 2. Given "junction" network $I^{\alpha} = Y^{\alpha\beta}E_{\beta}$.

desired to find the final equations $\mathbf{E}=\mathbf{ZI}$ of the physical system.

(2) Establish the equivalent circuit (electric circuit model) of the physical system.

(3) Divide the network into several component parts and establish the $\mathbf{I}=\mathbf{YE}$ equations of each component.

(4) Solve for the inverse equations of each part, namely for $\mathbf{E}=\mathbf{ZI}$, in any manner. (By actually inverting each $\mathbf{Y}$, or by a tensorial method suggested here, or by measurements on a dc analyzer, etc.)

(5) Establish a new "inverse" equivalent circuit for each solution.

(6) Interconnect the final $\mathbf{Z}$ matrix of each component part into the exact final $\mathbf{Z}$ matrix of the entire system by a series of transformations dictated by the interconnection of the new "inverse" equivalent circuits.

(7) Eliminate or solve for the constraints appearing at the interconnections.

The new method is definitely not identical with the inverse calculation of a large matrix by submatrices. Even with a ten or a hundred subdivisions of the resultant system, there are only two extra matrix multiplications to be performed with the new method, instead of the hundreds or thousands of matrix multiplications required by the method of submatrices.

The interconnection of a large number of subdivisions may be performed in several steps in succession, to reduce the size of the necessary additional inverse matrices. By judicious selection of the subdivisions even these extra inverse calculations may be greatly reduced.

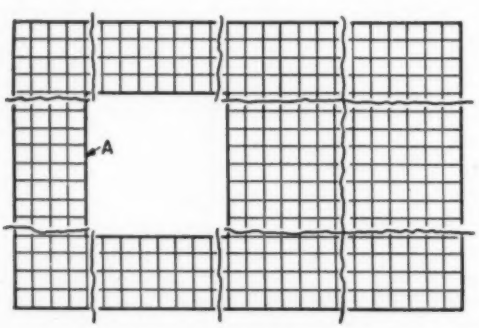


FIG. 3. Subdivision into eleven parts.

A labor-saving "skyscraper" method of interconnection will be shown in another publication.

Summary

The new method of interconnecting solutions of physical problems is based upon the known facts that with every physical phenomenon there are associated far more variables and equations than are customarily given. The new method does not ignore, but actually establishes these superfluous variables and equations by the artifice of setting up an equivalent electric circuit model and viewing the latter not as a "mesh" nor as a "junction," but as an "orthogonal" network. Afterward, these superfluous variables and equations are put to work to simplify the solution of the original set of equations by a new theory of "interconnection."

The device of analyzing in terms of an electrical network enables one to introduce the superfluous variables and their equations at the precise time when they are actually needed. The electric circuit model also supplies a ready-made "schedule" for the mathematical operations to be performed by the digital computer, that is hard to compete with in flexibility or versatility. Odd-shaped boundaries, physical systems with partly field and partly circuit properties (e.g., elastic field structures with reinforcements, transformer coils surrounded by their fields) cause no difficulties in model representation, even though their mathematical description runs into great difficulties.

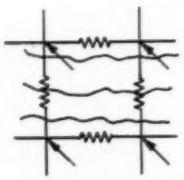


FIG. 4. Some possible cuts of a mesh.

Electric circuit models also facilitate the introduction of new types of iterative methods of solution. The circuit impedances may be "relaxed" independently of the repetitive steps, leading thereby to faster and definitive convergence. The process of tearing itself of course also speeds up the convergence of both conventional and new types of iterations.

SOLUTION OF THE SUBDIVISIONS

The Smallest Subdivision

It will be assumed that any available method may be used to establish the desired equations of the smallest subdivision, depending on the availability of a dc analyzer, or a digital computer, or of only an ordinary Marchant. Some of these possible methods for the problem under investigation will be reviewed shortly to establish a common starting point.

The main contribution of this paper will consist in interconnecting the solutions of the subdivisions in succession to any desired size.

The Problem

Let the planar resistance network of Fig. 2 (with 336 small meshes and one large mesh) correspond to Poisson's equation Divy grad$E=I$, where the components of matrix **y** are arbitrary functions of two variables. The potentials E are to be measured from point g (ground). The problem is to find the E of all junctions for any arbitrary set of given I's.

Let the network be subdivided into eleven parts, as shown in Fig. 3. If rectangular reference axes are assumed, there are only three different types of subdivisions that have to be solved, and not eleven.

The subdivision of a mesh is shown in Fig. 4. It is possible to leave either the whole, or part of the admittances **y** on either side of the cut. Strictly speaking the cut ought to be made in such a manner that no extra force and constraint appear, so that the resultant **C** would be square. In order to simplify the visualization, that requirement will be overlooked in the present simple problem, and an extra I is introduced at each side of the cut. After interconnection the extra I's will be dropped.

The equivalent circuit of subdivision A is shown in detail in Fig. 5(b). In each subdivision an arbitrary junction is assumed as ground.

Standard Solution of a Subdivision

The problem connected with each subdivision is to establish the **Z** matrix of its junction-pairs (not of the meshes). That is the *inverse* of the **Y** matrix of the junction equations **I=YE** need be established.

If a digital computer is available, then the **I=YE** equations are inverted to **E=ZI** by some iterative procedure, by an inverse calculation, by a relaxation method, etc.

If a dc calculating board is available having at least as many resistances as the subdivision, then the **Z** matrix of the junction-pairs may be established outright by the following measurements:

(1) Impress a unit current between one of the junction points and the ground.

(2) Measure the differences of potential appearing between the ground and each of the junction points.

(3) Repeat steps (1) and (2) for each junction point.

The measured differences of potential represent the desired **Z** matrix of the junction-pairs.

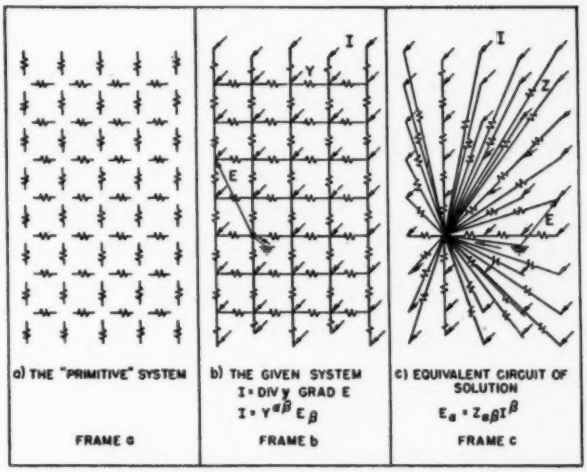


FIG. 5. Solution of one of the subdivisions.

Tensorial Solution of a Subdivision

The "tensorial" method establishes the solution of each component subdivision without finding any inverse matrices (that is using only elimination of superfluous variables). The same method can of course be employed for the solution of an entire system in one step, if the available digital computer can manipulate the large matrix that occurs.

Let subdivision A of Fig. 5(b) be solved by a tensorial method. The network has 45 junction points, hence 44 junction-pairs, also 24 meshes. The total number of resistances is 68 (=24+44). Although, it is possible to leave out temporarily the ten resistances at the two cuts, that labor-saving device is not made use of now. Three steps are made (TAN, page 98):

(a) All interconnections between the resistances of Fig. 5(b) are removed, leaving the "primitive" system of Fig. 5(a), to be called frame a. The equations of the 68 isolated coils can be written outright as $\mathbf{E}_a=\mathbf{Z}_{aa}\mathbf{I}^a$, where $\mathbf{Z}_{aa}$ has only diagonal components.

(b) Interconnect the "primitive" network of Fig. 5(a) (frame a) into the network of Fig. 5(b) (frame b) by a square (nonsingular) matrix of transformation $\mathbf{C}_b{}^a$ with 68 rows and columns.

To save space, the simpler system of Fig. 6 is used as an illustration for $\mathbf{C}_b{}^a$. The system has 10 resistances, 2 meshes, 9 junctions, hence 8 junction-pairs. By comparing the "old" and "new" currents flowing in each coil, the following relations $\mathbf{I}^a=\mathbf{C}_b{}^a\mathbf{I}^b$ exist:

$$
\begin{array}{ll}
I^1=I^{1'} & \\
I^2=I^{2'} & \\
I^3=I^{3'} & \\
I^4=I^{1'}+I^{4'}-i^a & \\
I^5=I^{2'}+I^{5'}+i^a-i^b & \mathbf{C}_b{}^a= \\
I^6=I^{3'}+I^{6'}+i^b & \\
I^7=i^a & \\
I^8=i^b & \\
I^9=I^{1'}+I^{4'}+I^{7'}-i^a & \\
I^{10}=I^{1'}+I^{2'}+I^{4'}+I^{5'}+I^{7'}+I^{8'}-i^b &
\end{array}
\begin{array}{c|cccccccc|cc}
 & 1' & 2' & 3' & 4' & 5' & 6' & 7' & 8' & a & b \\
\hline
1 & 1 & & & & & & & & & \\
2 & & 1 & & & & & & & & \\
3 & & & 1 & & & & & & & \\
4 & 1 & & & 1 & & & & & -1 & \\
5 & & 1 & & & 1 & & & & 1 & -1 \\
6 & & & 1 & & & 1 & & & & \\
7 & & & & & & & & & 1 & \\
8 & & & & & & & & & & 1 \\
9 & 1 & & & 1 & & & 1 & & -1 & \\
10 & 1 & 1 & & 1 & 1 & & 1 & 1 & & -1
\end{array}
\qquad (5)
$$

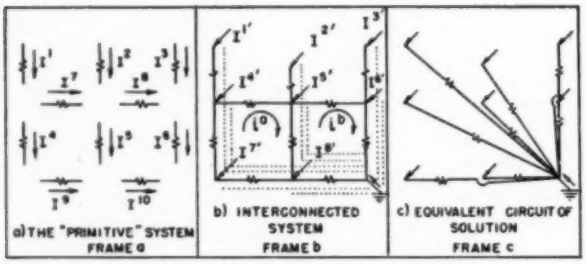


FIG. 6. Simple example of an "orthogonal" network.

The transformation matrix $\mathbf{C}_b{}^a$ splits by a vertical line into two components

$$\mathbf{C}_b{}^a = \mathbf{C}_j + \mathbf{C}_m.$$

If the network of Fig. 6(b) had been analyzed as a "mesh" network, only then its matrix of transformation (or "connection" tensor), changing the primitive system of Fig. 6(a) into the mesh network of Fig. 6(b), would be $\mathbf{C}_m$.

It should be mentioned that if no I are impressed at some of the junctions, then the corresponding columns of $\mathbf{C}_b{}^a$ are omitted.

The impedance matrix of frame b is found by $\mathbf{Z}_{bb} = (\mathbf{C}_b{}^a)_t \mathbf{Z}_{aa} \mathbf{C}_b{}^a$. (This formula of transformation leaves $\mathbf{IE}$ invariant.) Subdividing $\mathbf{Z}_{bb}$ along the junction-pairs and meshes (TAN, page 243) the 68 equations of frame b may also be written as $\mathbf{E}_b = \mathbf{Z}_{bb}\mathbf{I}^b$ (see TAN, page 422, where $\mathbf{e}$ is now zero).

$$\begin{matrix} & [\mathbf{I}^j & \mathbf{i}^m] \\ \begin{bmatrix} \mathbf{E}_j \\ 0 \end{bmatrix} = & \begin{bmatrix} \mathbf{Z}_1 & \mathbf{Z}_2 \\ \mathbf{Z}_3 & \mathbf{Z}_4 \end{bmatrix} \end{matrix}. \tag{6}$$

(c) Since the voltage vector $\mathbf{e}$ impressed upon the meshes is zero, the unknown mesh currents $\mathbf{i}^m$ may be eliminated by a matrix of transformation,

$$\mathbf{C}_e{}^b = \begin{matrix} & \mathbf{j} & \\ \mathbf{j} & \mathbf{1} & \\ \mathbf{m} & -\mathbf{Z}_4{}^{-1} & \mathbf{Z}_3 \end{matrix}. \tag{7}$$

The transformation $(\mathbf{C}_e{}^b)_t \mathbf{Z}_{bb} \mathbf{C}_e{}^b$ gives the final required $\mathbf{Z}_{ee}$ of all junction-pairs.

The above transformation is equivalent to the following two matrix multiplications (TAN, page 244, Eq. 10.7):

$$\mathbf{Z}_{ee} = \mathbf{Z}_1 - \mathbf{Z}_2 \mathbf{Z}_4{}^{-1} \mathbf{Z}_3. \tag{8}$$

In either case the inverse of $\mathbf{Z}_4$ must be calculated having as many rows as there are meshes in the subdivision. It should however be noted that, at least in the present example, $\mathbf{Z}_4$ has less rows (24) than $\mathbf{Y}$ of the conventional method (44) hence the inverse calculation with the tensorial method of attack is shorter.

The elimination may be performed one constraint at a time. That procedure becomes especially advantageous with digital computers. The elimination may be performed also by subdividing the axes to be eliminated into several sets and eliminating a set of variables at a time.

Summary of the Tensorial Work

It should be noted that the above tensorial method consists of further subdividing each subdivision, until the ultimate sub-subdivision is an isolated coil, frame **a**. Of course frame a—the "primitive" network—can be solved *outright* by simply writing out the desired solutions $\mathbf{E}_a = \mathbf{Z}_{aa}\mathbf{I}^a$. Then the solutions of each coil are interconnected into $\mathbf{E}_b = \mathbf{Z}_{bb}\mathbf{I}^b$. Finally all meshes introduced by the interconnection are eliminated.

The interconnection of solutions is rather simple in the present case since the connection tensor $\mathbf{C}_b{}^a$ of the primitive system can be written down by inspection only. In interconnecting the solutions of the subdivisions, however, the connection tensor will have to be established in four steps instead of one step.

The "Inverse" Equivalent Circuit of the Solution

No matter what method is used to find the solution $\mathbf{E} = \mathbf{ZI}$ of a subdivision, the resulting equations may be represented by an all-junction network with 44 electrically isolated coils (or an "n-winding transformer") having electromagnetic couplings between each of the 44 coils, as shown in Figs. 5(c) and 6(c). If so desired, it is possible to interconnect all coils at the ground point, thereby reducing the system from 44 electrically isolated networks to one single network. The number of junction-pairs is the same in either case (TAN, page 75, Eq. 3.1). The network of Fig. 5(c) is merely a simplified form of the original network of Fig. 5(b) with all meshes eliminated. Also now impedances Z are used to define the coils, instead of admittances Y.

The entire system of Fig. 3 can be reduced to eleven isolated "inverse" networks, having no mutual impedances between them. (Only three different calculations need to be made for the circuit of Fig. 3.) The tensorial reduction of Fig. 3 to the present case required the elimination of all uncut meshes of Fig. 3 (namely of $337 - 59 = 278$ meshes).

The interconnection of the eleven subdivisions will entail only the elimination of the remaining 59 meshes of the original network of Fig. 2. That elimination will require, however, a transformation of the remaining meshes to another set of meshes by a series of four transformations, as will be shown presently.

The Proof of "Orthogonality"

At this point, it will be proved that the configuration space, representing the meshes alone of any network, is orthogonal (normal) to the configuration space representing the junction-pairs alone of the same network.

Using the simple network of Fig. 6 as an illustration, let the admittance matrix of the primitive network of Fig. 6(a) be $\mathbf{Y}^{aa} = \mathbf{Z}_{aa}{}^{-1}$. To find $\mathbf{Y}^{bb}$ of the interconnected system of Fig. 6(b), it is necessary to establish a relation between the "old" and the "new" junction-pair potentials as $\mathbf{E}_a = \mathbf{A}_a{}^b \mathbf{E}_b$ (TAN, page 359). The

resulting nonsingular (square) connection tensor is

$$\mathbf{A}_a{}^b = \begin{array}{l|cccccccccc} & 1' & 2' & 3' & 4' & 5' & 6' & 7' & 8' & a & b \\ E_1=E_1'-E_4' \quad 1 & 1 & & & -1 & & & & & & \\ E_2=E_2'-E_5' \quad 2 & & 1 & & & -1 & & & & & \\ E_3=E_3'-E_6' \quad 3 & & & 1 & & & -1 & & & & \\ E_4=E_4'-E_7' \quad 4 & & & & 1 & & & -1 & & & \\ E_5=E_5'-E_8' \quad 5 & & & & & 1 & & & -1 & & \\ E_6=E_6' \quad 6 & & & & & & 1 & & & & \\ E_7=E_4'-E_5'+e_a \quad 7 & & & & 1 & -1 & & & & 1 & \\ E_8=E_5'-E_6'+e_b \quad 8 & & & & & 1 & -1 & & & & 1 \\ E_9=E_7'-E_8' \quad 9 & & & & & & & 1 & -1 & & \\ E_{10}=E_8' \quad 10 & & & & & & & & 1 & & \end{array}. \tag{9}$$

Since

$$\mathbf{Y}^{bb} = (\mathbf{A}_a{}^b)_t \mathbf{Y}^{aa} \mathbf{A}_a{}^b,$$

and

$$\mathbf{Z}_{bb} = (\mathbf{C}_b{}^a)_t \mathbf{Z}_{aa} \mathbf{C}_b{}^a, \tag{10}$$

also

$$\mathbf{Y}^{aa} = \mathbf{Z}_{aa}{}^{-1} \mid \mathbf{Y}^{bb} = \mathbf{Z}_{bb}{}^{-1}. \tag{11}$$

Taking the inverse of the second equation,

$$\mathbf{Y}^{bb} = (\mathbf{C}_b{}^a)^{-1} \mathbf{Y}^{aa} (\mathbf{C}_b{}^a)_t{}^{-1}. \tag{12}$$

Comparing the first and last equations

$$\mathbf{A}_a{}^b = (\mathbf{C}_b{}^a)_t{}^{-1},$$

or

$$\mathbf{C}_b{}^a{}_t \mathbf{A}_a{}^b = \mathbf{1}. \tag{13}$$

That is the two nonsingular connection tensors **C** and **A** are inverse to each other.

$\mathbf{A}_a{}^b$ also splits into two components

$$\mathbf{A}^b = \mathbf{A}_j + \mathbf{A}_m = \begin{matrix} \mathbf{j} & \mathbf{m} \end{matrix} \begin{bmatrix} \mathbf{A}_j & \mathbf{A}_m \end{bmatrix}.$$

If the network of Fig. 6(b) had been analyzed as a "junction" network only, then its matrix of transformation would be $\mathbf{A}_j$. However, since **C** and **A** are inverse to each other

$$(\mathbf{C}_b{}^a)_t (\mathbf{A}_a{}^b) = \mathbf{1} = \begin{bmatrix} \mathbf{C}_{jt} \\ \mathbf{C}_{mt} \end{bmatrix} \begin{bmatrix} \mathbf{A}_j & \mathbf{A}_m \end{bmatrix}$$

$$= \begin{matrix} \mathbf{j} \\ \mathbf{m} \end{matrix} \begin{bmatrix} \mathbf{C}_{jt}\mathbf{A}_j & \mathbf{C}_{jt}\mathbf{A}_m \\ \mathbf{C}_{mt}\mathbf{A}_j & \mathbf{C}_{mt}\mathbf{A}_m \end{bmatrix} = \begin{matrix} \mathbf{j} \\ \mathbf{m} \end{matrix} \begin{bmatrix} \mathbf{j} & \mathbf{m} \\ 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{14}$$

It follows that if the connection tensor $\mathbf{C}_m$ of a purely mesh network point of view is multiplied by the connection tensor $\mathbf{A}_j$ of a purely junction network point of view, their product is always zero,

$$\mathbf{C}_{mt} \mathbf{A}_j = \mathbf{0}. \tag{15}$$

$$\begin{array}{l|cccccccccc} & 1' & 2' & 3' & 4' & 5' & 6' & 7' & 8' \\ 1 & 1 & & & -1 & & & & \\ 2 & & 1 & & & -1 & & & \\ 3 & & & 1 & & & -1 & & \\ 4 & & & & & & & -1 & \\ 5 & & & & & 1 & & & -1 \\ 6 & & & & & & 1 & & \\ 7 & & & & 1 & -1 & & & \\ 8 & & & & & 1 & -1 & & \\ 9 & & & & & & & 1 & -1 \\ 10 & & & & & & & & 1 \end{array} = 0.$$

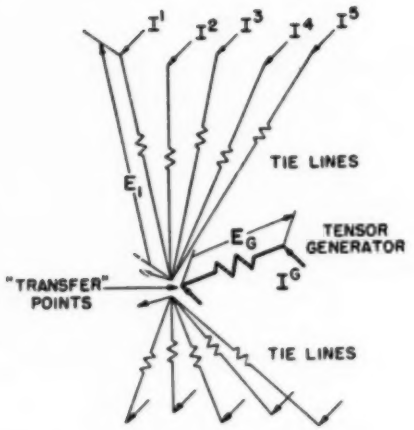
$$\begin{array}{l|cccccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ a & & & & -1 & 1 & & 1 & & -1 & \\ b & & & & & -1 & 1 & & 1 & & -1 \end{array}$$

This result is a generalization of the well-known fact that the product of two orthogonal vectors is zero. Here a large number of vectors are mutually perpendicular. Similarly

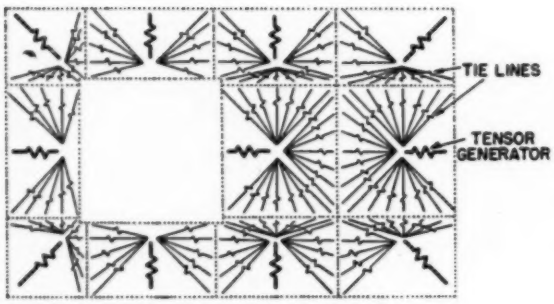
$$\mathbf{C}_{jt} \mathbf{A}_m = \mathbf{0}. \tag{16}$$

Hence, the name "orthogonal" network for one containing both meshes and junction-pairs. Or, better expressed, every electrical network (or any other network) without exception is an orthogonal network by its very existence. However, depending upon the limitations of the problem at hand, attention is restricted usually either to the mesh or to the junction-pair portions of the entire system.

Because of this orthogonality, all n-coil networks may be transformed into each other by a set of nonsingular connection matrices $\mathbf{C}_b{}^a$ that form a "group." The above $\mathbf{C}_b{}^a$ is one element of that group. It transforms the 10-coil network of Fig. 6(a) to the 10-coil network of Fig. 6(b). From the same ten coils a large variety of networks may be constructed, all forming a group.



(a)



(b)

FIG. 7. Equivalent circuit of the *component* solutions. (a) Generalized "force" variables E_α and I^α (subdivision A). (b) The "primitive" system.

INTERCONNECTION OF SOLUTIONS

Transmission System Analogies

The solution $\mathbf{E}_c=\mathbf{Z}_{cc}\mathbf{I}^c$ of each subdivision is represented by the n-winding transformer of Fig. 5(c). It should be expressly noted that the set of equations give no clue whatever whether the end-points of the n coils should all be isolated, or connected to one junction point, or to several junction points. Each of these possibilities gives the correct number of junction-pairs (and zero meshes). The nature of the interconnection to follow will dictate the correct number of junction-points.

In order to give a physical picture of the steps to follow, it will be assumed that each subdivision represents an electric power transmission system belonging to some independent company. The junction points will be divided into three sets:

(1) Each junction point that is not the result of a cut mesh, represents a "generator" (a power plant) that feeds current I into the transmission system.

(2) The opened lines of the cut meshes represent "tie-lines" that carry currents (power) from the given system into the neighboring system (or vice versa). The tie lines of each subdivision will be grouped into as many sets, as there are neighboring subdivisions.

(3) The ground junction carries the "load" current. Although in a transmission system there are a large number of loads, they are often replaced by *one* "equivalent" (or "hypothetical") load for analytical purposes (TAIT-I, Eq. 2).

Tensor Networks

In tensor analysis a large number of physical entities are replaced by one symbol, represented by a boldface letter as $\mathbf{Z}$. Similarly in the equivalent circuit the corresponding large number of coils may be replaced by one coil, represented by a boldface sketch of a coil (TAN, page 480 and TAIT-III, Chapter V).

For the time being all generator coils are replaced by one tensor coil $\mathbf{Z}_g$ as shown in Fig. 7(a), while the tie-lines are still shown by scalar coils.

Splitting the Load Point (Frame 1)

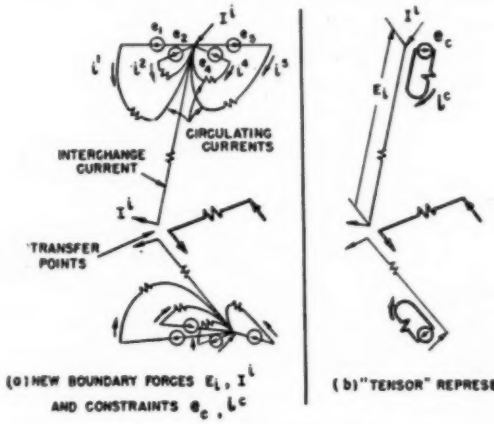
It will also be assumed (to be justified by the step to follow) that the common load point should be split into several points (points of transfer, TAIT-IV, Chapter II). In particular:

(1) All generators should be tied to a common point.

(2) All tie-lines leading to each neighboring subdivision should also be tied to a common point [Fig. 7(a)].

Physically these common junctions all lie at the same junction point, but they may assume (for the time being) different absolute potentials.

The network of Fig. 7(a) possesses the same equations $\mathbf{E}_c=\mathbf{Z}_{cc}\mathbf{I}^c$ as the network of Fig. 5(c) does. Hence no



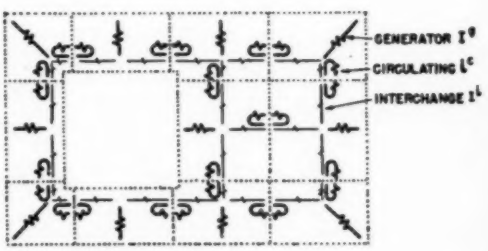


FIG. 8. Introduction of unknown boundary forces and constraints.

transformation of reference frames occurs, only a different physical representation of the same equations. Figure 7(b) of the entire system will be the starting point of interconnecting the solutions $\mathbf{E}_c=\mathbf{Z}_{cc}\mathbf{I}^c$. It will be therefore denoted as frame 1 and the equations as $\mathbf{E}_1=\mathbf{Z}_{11}\mathbf{I}^1$.

Introduction of Mesh Variables (Frame 2)

Frame 1 of each subdivision is an all-junction network containing no meshes. If two neighboring subdivisions are interconnected at their tie-lines, the interconnected system contains also meshes, in addition to junction-pairs. In order that the number of new variables should correspond to the correct degrees of freedom of the new network, a set of mesh variables will be introduced before the interconnection (TAIT-III, Chapter II).

In each set of tie-lines (leading to each neighboring subdivision) let one tie-line be selected arbitrarily and called "interchange" (or "exchange") tie-line. Let it be assumed that this particular tie-line carries the *sum* of all tie-line currents into the neighboring subdivision. The currents entering through the other tie-lines will then merely circulate through the subdivision and depart at the interchange tie-line. The corresponding nonsingular transformation of the upper set of Fig. 8(a) is (TAIT-III, Eq. 3)

$$\begin{aligned} I^1&=i^{1'}\\ I^2&=i^{2'}\\ I^3&=-I^{1'}-I^{2'}-I^{4'}-I^{5'}+I^i\\ I^4&=i^{4'}\\ I^5&=i^{5'} \end{aligned} \qquad \mathbf{C}_2{}^1=\begin{array}{c} \\ 1\\2\\3\\4\\5 \end{array}\begin{array}{c} \begin{array}{ccccc} 1' & 2' & 4' & 5' & \mathbf{i} \end{array}\\ \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ -1 & -1 & -1 & -1 & 1 \\ & & 1 & & \\ & & & 1 & \end{bmatrix} \end{array}. \tag{17}$$

The large number of circulating meshes of each set of tie-lines may be replaced by a tensor mesh, as shown in Fig. 8(b) with an impressed voltage **e**. The tensor network of frame 2 of all the subdivisions is shown in Fig. 8(c). Since the circulating coil is tied only at *one* point to the interchange coil, the two coils may be shown electrically isolated. However, within each subdivision mutual inductances exist between all the coils. (No mutual impedances exist between the subdivisions themselves.) It should be noted that the interchange and generator coils still do not have a common junction point.

It should also be emphasized that the order of the transformations $\mathbf{C}_2{}^1$ and $\mathbf{C}_3{}^2$ may be interchanged. The neighboring tie-lines may at first be interconnected and then only are the interchange and circulating paths introduced. As this reverse order would introduce (temporarily only) redundant variables, the given order is preferred. (Of course the two transformations may also be performed in one step as $\mathbf{C}_3$.)[1]

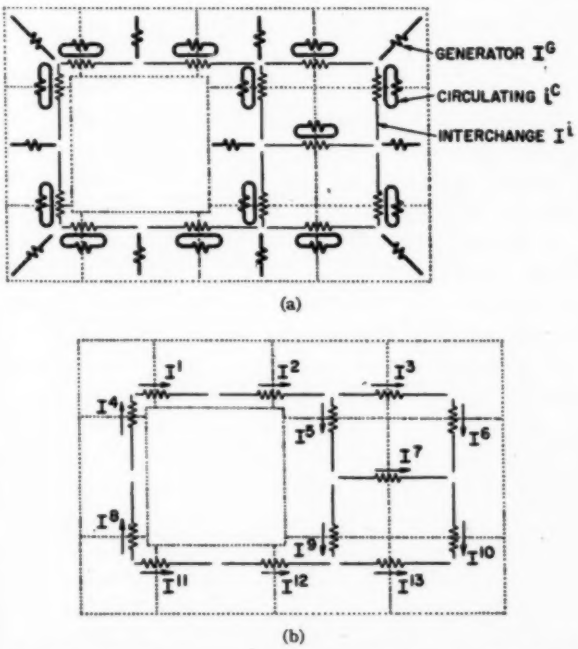


FIG. 9. The interconnection of boundaries. (a) The "primitive interconnected system. (b) The "primitive" interchange network.

The "Primitive" Interconnected System (Frame 3)

Let next all circulating meshes and interchange coils of neighboring subdivisions be interconnected in series, as shown in Fig. 9(a). Two neighboring circulating (tensor) meshes become one mesh and two interchange radial coils become one. If the five upper tie-lines of Fig. 8(a) (to be denoted by primes) are interconnected in series with the analogous five lower tie-lines of the neighboring division (to be denoted by double primes) the connection tensor is (TAIT-IV, Eq. 1)

$$\mathbf{C}_3{}^2=\begin{array}{c} \\ 1'\\2'\\4'\\5'\\3'\\1''\\2''\\4''\\5''\\3'' \end{array}\begin{array}{c} \begin{array}{ccccc|c} 1 & 2 & 4 & 5 & 3 & \end{array}\\ \left[\begin{array}{ccccc|c} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & & 1 \\ -1 & & & & & \\ & -1 & & & & \\ & & -1 & & & \\ & & & -1 & & \\ & & & & -1 & \end{array}\right] \end{array} = \begin{array}{c} \\ m'\\j'\\m''\\j'' \end{array}\begin{array}{c} \begin{array}{cc} m & j \end{array}\\ \begin{bmatrix} 1 & \\ & 1 \\ -1 & \\ & -1 \end{bmatrix} \end{array}. \tag{18}$$

The generators are left undisturbed.

As was mentioned before, the cut between these two divisions could have been made in such a manner that

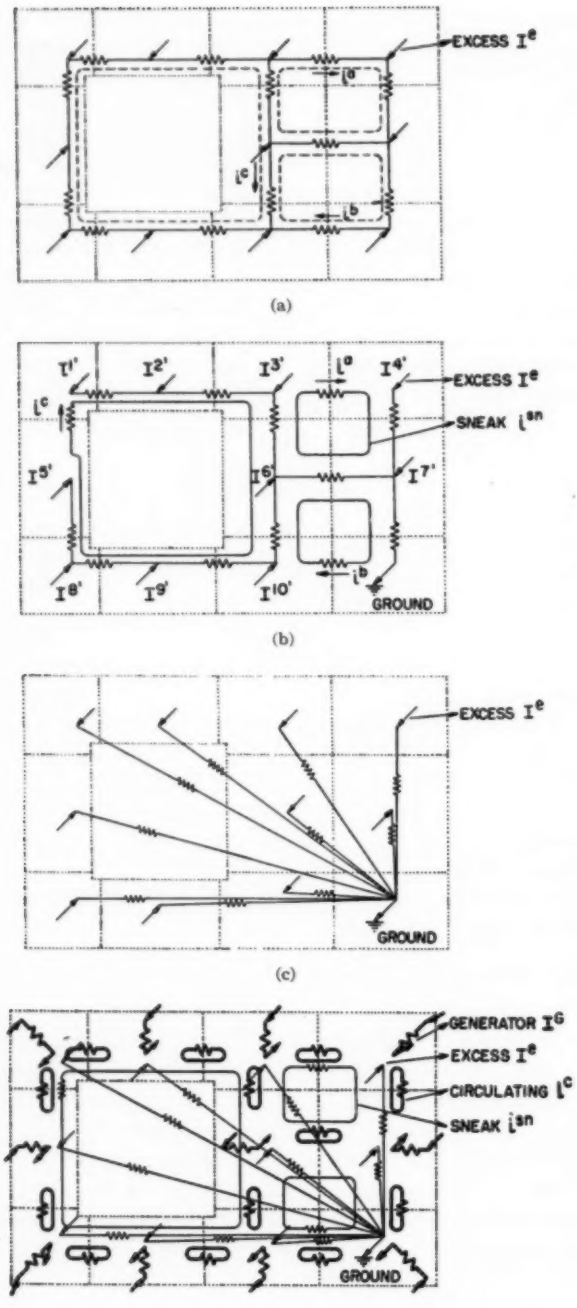


FIG. 10. The interconnection of "interchange" network. (a) The "interconnected" interchange network. (b) The "tree" and "loops" (the "radial" and "sneak" paths). (c) Another representation of the "radial" paths. (d) The resultant system of frame 4.

the above C_3^2 would have been nonsingular (square). For the sake of better visualization more variables were introduced at the cut than were absolutely necessary. But now the superfluous variables drop out.

It should be expressly noted that all generators and interchange coils are still isolated at the load points. The effect of the series interconnection is to reduce to one-half the number of circulating meshes and interchange paths, leaving the latter still isolated from the generators. Hence Fig. 9(a) may be called the "primitive" interconnected system.

The Interchange Network (Frame 4)

Next attention will be turned to interconnecting the interchange coils only. The "primitive" interchange network of Fig. 9(a) before interconnection is shown again in Fig. 9(b), in which each of the thirteen coils is isolated, as in any true "primitive" network (such as frame *a* of Fig. 5(a). Each coil carries the sum of all tie-line currents (the "interchange" current $\mathbf{I}^i$) crossing the boundaries between two subdivisions.

Let now the thirteen interchange coils of Fig. 9(b) be interconnected at their junction-points into frame 4, as shown in Fig. 10(a). The interconnected system has as many junction points as there are subdivisions, namely, eleven. Each current I^e impressed at the junctions represents the sum of all interchange currents $\mathbf{I}^i$ (the sum of all sum-currents) leaving a subdivision. That is each junction current represents the "excess" current I^e of each subdivision, namely, the difference between all its generated currents $\mathbf{I}^G$ and load current I^L. It should be noted that the sum of all excess currents I^e in the entire system is zero.

The interconnected interchange network of Fig. 10(a) has thirteen coils and eleven junctions. Hence it has ten junction-pairs and three meshes. That is the total system happened to be so subdivided that the subdivisions (as represented by the interchange network) form three closed loops. It should also be noted that the number of junction-pairs is $d-1$, if d is the number of subdivisions (or junctions).

"Basic" Meshes and Junction-Pairs

Although it is theoretically possible to assume the meshes and junction-pairs in an arbitrary manner, for the present purpose a so-called "basic" set will be assumed as follows [Fig. 10(b)] (TAIT-IV, Chapter III):

(1) Remove as many interchange coils as there are meshes (three), observing that the remaining coils still pass through all junctions. The resulting all-junction network is called in topology a "tree" (TAN, page 213).

(2) Assume that all junction-pairs pass through the ground point g of the entire system, as shown in Fig. 10(c).

(3) Each of the removed interchange coils forms a separate closed mesh, passing through the coils of the tree also. (In a transmission system the so-called "sneak" powers circulate in these meshes between the various companies.) It should be noted that these sneak currents $\mathbf{i}^{sn}$ circulate between three or more subdivisions, while the previous circulating currents $\mathbf{i}^c$ circulate always between two neighboring subdivisions.

The matrix of transformation of Figs. 10(a) or 10(b)

is a nonsingular matrix (TAN, page 419, Eq. 16.6),

$$C_4^3 = \begin{array}{c|cccccccccc|ccc} & 1' & 2' & 3' & 4' & 5' & 6' & 7' & 8' & 9' & 10' & a & b & c \\ 1 & 1 & & & & & & & & & & & & 1 \\ 2 & 1 & 1 & & & & & & & & & & & 1 \\ 3 & & & & & & & & & & & 1 & & \\ 4 & & & & & & & & & & & & & 1 \\ 5 & 1 & 1 & 1 & & & & & & & & -1 & & 1 \\ 6 & & & & 1 & & & & & & & 1 & & \\ 7 & 1 & 1 & 1 & & & 1 & & 1 & 1 & 1 & -1 & 1 & \\ 8 & & & & & -1 & & & & & & & & 1 \\ 9 & & & & & -1 & & & -1 & -1 & -1 & & -1 & 1 \\ 10 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & & 1 & \\ 11 & & & & & 1 & & & 1 & & & & & -1 \\ 12 & & & & & 1 & & & 1 & 1 & & & & -1 \\ 13 & & & & & & & & & & & & -1 & \end{array} . \qquad (19)$$

The equation of the resultant system of Fig. 10(d) is $E_4 = Z_{44}I^4$, where I^4 contains (1) generator currents I^G, (2) circulating currents i^c, (3) sneak (or loop) currents i^{sn}, and (4) excess currents I^e.

All excess currents I^e flow into the system ground g, but the generator currents I^G do not. Each generator still delivers its current into the original ground of each subdivision.

Appearance of the Load Currents (Frame 5)

In order that each generator should deliver its current into the system ground g, it is necessary to replace all excess currents I^e by the actual "load" currents I^L existing at the ground point of each subdivision. Since at each load point (TAIT-IV, Eq. 5)

$$I^e = \Sigma I^G - I^L = 1_v I^G - I^L,$$

(where 1_v is a horizontal vector with all its component unity) the matrix of transformation for the eleven subdivisions is

$$C_5^4 = \begin{array}{c|cccc|cccc} & G_1 & G_2 & \cdots & G_{11} & L_1 & L_2 & \cdots & L_{11} \\ G_1 & 1 & & & & & & & \\ G_2 & & 1 & & & & & & \\ \cdots & & & \cdots & & & & & \\ G_{11} & & & & 1 & & & & \\ e_1 & 1_v & & & & -1 & & & \\ e_2 & & 1 & & & & -1 & & \\ \cdots & & & \cdots & & & & \cdots & \\ e_{11} & & & & 1_v & & & & -1 \end{array} . \qquad (20)$$

The equation of the resultant system of Fig. 11 is $E_5 = Z_{55}I^5$. The currents $\mathbf{I}$ represent all the junction currents flowing into the system ground, plus the unknown circulating and sneak currents $\mathbf{i}$. The voltages E_5 represent the differences of potential appearing between each junction point and the system ground point g. The meshes now have no impressed voltage $\mathbf{e}$.

The interconnection of the component solutions has been performed.

Elimination of Meshes (Frame 6)

The final step consists in eliminating the mesh currents by the transformation shown in Eq. (7) or by the formula of Eq. (8). It requires the calculation of the inverse of a matrix having as many rows and columns as there are circulating and sneak meshes. Again one mesh current at a time may be eliminated.

It should be especially noted that there are as many circulating and sneak meshes to eliminate as there exist cut plus disappeared meshes in Fig. 3. (There were 58 meshes cut in Fig. 3, also the large mesh has disappeared because of the cuts. The number of circulating meshes is 56, and the number of sneak meshes is 3. Their total is 59.) If the digital computer available is not able to invert this last matrix, then the interconnection of subdivisions may be performed in two, or even more series of steps.

The resultant $E_6 = Z_{66}I^6$ equations of Fig. 12 are identical with those found when the entire system equations $I^6 = Y^{66}E_6$ are inverted all at once.

The resulting system of Fig. 12 is ready to be interconnected with other systems to form a partial solution of a still larger system.

Résumé of Steps

It should be noted that the first four transformation matrices performing the actual interconnection of the component solutions include only ± 1 and 0. Hence if

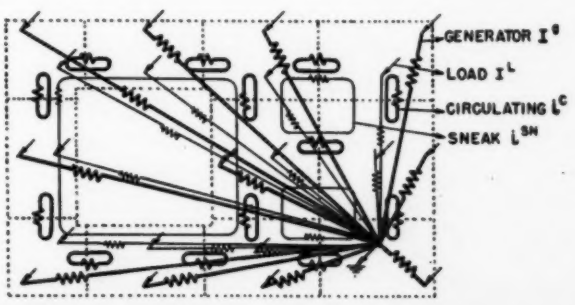


FIG. 11. Replacing the "excess" currents by "load" currents.

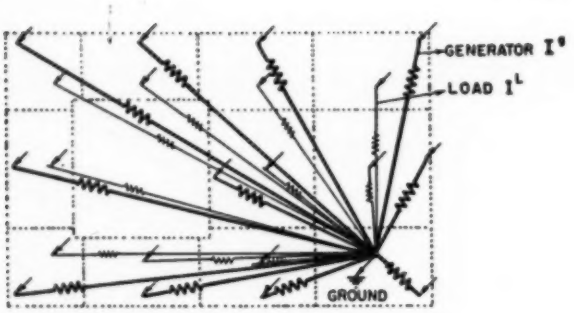


FIG. 12. Equivalent circuit of the resultant solution.

each of them has been established, their resultant is

$$\mathbf{C}_5^1=\mathbf{C}_2^1\mathbf{C}_3^2\mathbf{C}_4^3\mathbf{C}_5^4, \tag{21}$$

where $\mathbf{C}_5^1$ still contains only ± 1 and 0. If now the $\mathbf{Z}_{11}$ of each subdivision is available, they are interconnected into one resultant system by

$$\mathbf{Z}_{55}=\mathbf{C}_{5t}^1\mathbf{Z}_{11}\mathbf{C}_5^1. \tag{22}$$

The product merely *adds* up the components of $\mathbf{Z}_{11}$ without multiplication. $\mathbf{Z}_{55}$ is subdivided along the junction-pairs and meshes as

$$\mathbf{Z}_{55}=\begin{matrix}\\ \mathbf{j}\\ \mathbf{m}\end{matrix}\begin{matrix}\begin{matrix}\mathbf{j} & \mathbf{m}\end{matrix}\\ \begin{bmatrix}\mathbf{Z}_1 & \mathbf{Z}_2\\ \mathbf{Z}_3 & \mathbf{Z}_4\end{bmatrix}\end{matrix}. \tag{23}$$

The mesh currents (the superfluous constraints) are eliminated in the quickest manner by first finding the inverse of $\mathbf{Z}_4$ and then finding

$$\mathbf{Z}_{66}=\mathbf{Z}_1-\mathbf{Z}_2\mathbf{Z}_4^{-1}\mathbf{Z}_3. \tag{24}$$

(The use of $\mathbf{C}_6^5$ is only for analytical purposes, as it also calculates two other products that cancel.)

Hence, the entire interconnection involves one smaller inverse calculation, two matrix multiplications and some additions. In the resulting equation $\mathbf{E}_6=\mathbf{Z}_{66}\mathbf{I}^6$ the currents $\mathbf{I}^6$ impressed at the junctions still may assume arbitrary values.

Solving for the Constraints

Since the component matrices $\mathbf{Z}_1$, $\mathbf{Z}_2$, and $\mathbf{Z}_3$ are strongly diagonal with many zero off-diagonal elements, *it is preferable not to perform the elimination*, but to leave Eq. (24) in its given form. That is the solution of $\mathbf{I}=\mathbf{YE}$ is represented by the form

$$\mathbf{E}=(\mathbf{Z}_1-\mathbf{Z}_2\mathbf{Y}^4\mathbf{Z}_3)\mathbf{I}. \tag{25}$$

For any given numerical $\mathbf{I}$ the substitutions indicated in Eq. (25) are to be performed in succession.

The new form of solution is equivalent to solving the more extensive orthogonal equations (4)

$$\mathbf{i}+\mathbf{I}=\mathbf{Y}(\mathbf{E}+\mathbf{e}), \tag{26}$$

giving the more general orthogonal solution

$$\mathbf{e}+\mathbf{E}=\mathbf{Z}(\mathbf{I}+\mathbf{i}), \tag{27}$$

where $\mathbf{I}$ and $\mathbf{e}$ are known ($\mathbf{e}=\mathbf{0}$), while $\mathbf{E}$ and $\mathbf{i}$ are unknown. The intermediate product $\mathbf{Y}^4\mathbf{Z}_3\mathbf{I}$ gives the unknown constraints $\mathbf{i}$.

The "orthogonal" inverse matrix may be represented as

$$\mathbf{Z}_{66}'=\begin{bmatrix}\mathbf{Z}_1 & \mathbf{Z}_2\\ \mathbf{Z}_3 & \mathbf{Z}_4^{-1}\end{bmatrix}, \tag{28}$$

differing from (23) by the presence of the constraint inverse $\mathbf{Z}_4^{-1}=\mathbf{Y}^4$. The "orthogonal" inverse matrix is strongly diagonal and contains only a small fraction of the elements that exist in the conventional inverse matrix $\mathbf{Z}_{66}$. In the present example of about 400 equations and 11 subdivisions, the original smaller inverse $\mathbf{Z}_{66}$ contains $400^2=160\,000$ elements, while the much larger orthogonal inverse matrix $\mathbf{Z}_{66}'$ contains only about 30 000 elements. For every given $\mathbf{I}$ the calculation of $\mathbf{E}$ by Eq. (25) or (28) involves fewer than one-fifth of the conventional number of multiplications.

Thus another important consequence of the process of tearing is that it requires far less computation to solve simultaneously for both forces and constraints, than to solve for the forces alone. This conclusion only corroborates the conception of the author that the absence of constraints destroys the simplicity of a model (be it a physical or a mathematical model).

Of course, if the impressed $\mathbf{I}$ contains only a few components, then the number of elements in the final matrix $\mathbf{Z}_{66}'$ are still more reduced. If furthermore $\mathbf{E}$ is needed at a few points only, the reduction in $\mathbf{Z}_{66}'$ is even greater.

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 8 AUGUST, 1953

Contact and Rubbing of Flat Surfaces

J. F. ARCHARD
Research Laboratory, Associated Electrical Industries Limited, Aldermaston, Berkshire, England
(Received January 8, 1953)

The interpretation of certain phenomena occuring at nominally flat surfaces in stationary or sliding contact is dependent on the assumed distribution of the real area of contact between the surfaces. Since there is little direct evidence on which to base an estimate of this distribution, the approach used is to set up a simple model and compare the deduced theory (e.g., the deduced dependence of the experimental observables on the load) with the experimental evidence. The main conclusions are as follows. (a) The electrical contact resistance depends on the model used to represent the surfaces; the most realistic model is one in which increasing the load increases both the number and size of the contact areas. (b) In general, mechanical wear should also depend on the model. However, in wear experiments showing the simplest behavior, the wear rate is proportional to the load, and these results can be explained by assuming removal of lumps at contact areas formed by plastic deformation; moreover, this particular deduction is independent of the assumed model. This suggests that a basic assumption of previous theories, that increasing the load increases the number of contacts without affecting their average size, is redundant.

I. GENERAL INTRODUCTION

WHEN two nominally mating surfaces are brought together, they touch at the tips of the higher asperities and the total area of intimate contact is determined by the deformation of the material in these regions under the applied load. The real area of contact, which is, in general, a small part of the apparent area of contact, therefore comprises several individual areas, the number, size, and distance of separation of which will influence the results of certain types of experiment. For example, it has been recognized that these factors will affect the contact resistance or the local temperatures caused by frictional heating during sliding (Bowden and Tabor[1]; Holm[2]);* also the character and magnitude of the transfer of material between rubbing surfaces and their wear rate may show a similar dependence. In many experiments, in order to overcome the uncertainties as to the distribution of the real area of contact, crossed cylinders or a sphere pressed against a flat are used. It is then assumed that the total area of contact is confined to a single circular area and that the experiment represents the conditions occurring at the individual contacts between the mating surfaces. However, in order to apply the results of such experiments to the more general case, it is necessary to make some assumptions about the distribution of the real areas of contact and then to deduce the effect of this on the behavior of mating surfaces. This paper represents an attempt to tackle the problem.

There is little direct evidence on which to base any estimate of the sizes and distribution of the individual areas involved in the contact of mating surfaces. Therefore, the approach used is to set up a simple model to represent the surfaces and to compare the theoretical deductions with the results of experiments. The multiple contact condition will be considered in terms of the contact of two nominally flat surfaces. At the lightest loads, it is assumed that they touch over three small areas. As the load is increased, these areas increase in size; also some of the asperities, hitherto separated, approach and touch to form new areas of contact. Thus, it is assumed that an increase in the load increases both the number and size of the contact areas; at the same time it automatically follows that these contact areas have a size distribution. Using a model which incorporates these features, the relationships between the applied load and the contact area or electrical conductance have been deduced and are compared with the corresponding expressions for the single contact condition. Making further simple assumptions as to the nature of the sliding process, simple laws of mechanical wear have been deduced. These theoretical conclusions are discussed in the light of the experimental evidence.

[1] F. P. Bowden and D. Tabor, *The Friction and Lubrication of Solids* (Oxford University Press, London, 1950).
[2] R. Holm, *Electrical Contacts* (H. Gerbers, Stockholm, 1946).
* These two references will be referred to throughout this paper as B.T. and E.C.H., respectively.

II. CONTACT AREA AND CONDUCTANCE

1. Introduction

The radius a of the circular area of contact formed when a sphere of radius R is pressed against a flat surface under a load P is given by†

$$a=1\cdot1\left\{\tfrac{1}{2}PR\left(\frac{1}{E_1}+\frac{1}{E_2}\right)\right\}^{\frac{1}{3}}, \tag{1}$$

when the deformation is elastic, E_1 and E_2 being the Young's moduli of the materials of the contacting members, and

$$a=(P/\pi p_m)^{\frac{1}{2}}, \tag{2}$$

when the deformation is plastic, p_m being the flow pressure (assumed constant) of the deformable member.

The conductance G of such a contact is given by

$$G=2a\sigma, \tag{3}$$

† For details of Eq. (1) to (4) see B.T. and E.C.H.

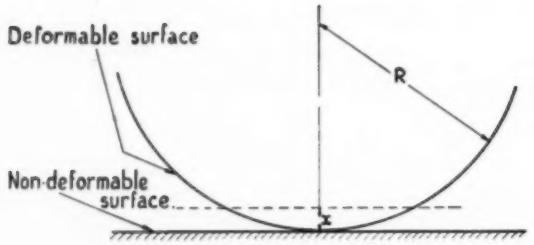


FIG. 1. Single area of contact model.

when the contact is purely metallic, the resistance of the contact being the "constriction resistance," and σ is the specific conductivity of the two materials (assumed equal). Also,

$$G=\pi a^2/\rho, \tag{4}$$

when the resistance of the contact is due to a film of resistance ρ per unit area, the "constriction resistance" being negligible.

Equations (1) to (4) will be used to calculate the contact area and conductance as a function of the load both for a single contact (representing a sphere pressed against a flat), and for the multiple areas of contact model (representing two flat surfaces in contact).

2. Single Area of Contact

The model assumed is shown in Fig. 1. It consists of a perfectly flat nondeformable surface and a deformable spherical surface of radius of curvature R. Changes in the area of contact A_1, the load P_1 carried by the contact, and the contact conductance G_1, are considered for a movement of the nondeformable surface through a distance x. A_1, P_1, and G_1 are given by the following general equations:

$$\left.\begin{aligned} A_1&=bx \\ P_1&=cx^p \\ G_1&=dx^m \end{aligned}\right\}, \tag{5}$$

where, using Eq. (1) to (4) above,

$$b=2\pi R,\ddagger$$

and

for plastic deformation $p=1$ and $c=2\pi Rp_m$,
for elastic deformation $p=\frac{3}{2}$ and $c=4.25ER^{\frac{1}{2}}$,
for a film resistance $m=1$ and $d=2\pi R/\rho$,

and

for a constriction resistance $m=\frac{1}{2}$ and $d=(8R)^{\frac{1}{2}}\sigma$.

These results give the following generalized relationships between the contact area or conductance and the load:

$$\left.\begin{aligned} A_1&=b(P_1/c)^{1/p} \\ G_1&=d(P_1/c)^{m/p} \end{aligned}\right\}. \tag{6}$$

‡ This may be incorrect by a factor of 2 for elastic deformation, but the subsequent deductions are not materially affected.

3. Multiple Areas of Contact

The model assumed is shown in Fig. 2. It consists of a perfectly flat nondeformable surface and a nominally flat deformable surface containing a large number of spherical asperities of equal radii of curvature R. It is assumed that the asperities are evenly distributed in depth (the x direction); i.e., there is one asperity at each of the x coordinates $x=0$, h, $2h$, $3h$ . . . etc., where $h\ll R$. These will be termed the zero, first, second, third . . . etc., levels, respectively. Thus, there are M asperities per unit depth, where

$$M=1/h.$$

Consider the total contact area A_2, load P_2, and conductance G_2 resulting from the movement of the nondeformable surface through a distance $x=Nh$ (i.e., such as to bring N asperities into contact). Then, if δA_r is the contribution to the total contact area A_2 from the asperity in the rth level, A_2 is the sum of N such areas, i.e.,

$$A_2=\sum_{r=0}^{r=N-1}\delta A_r=bh\sum_{r=1}^{r=N}r,$$

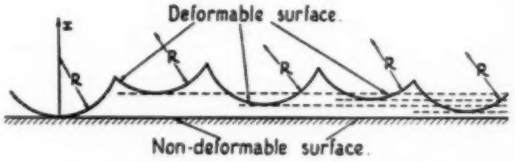


FIG. 2. Multiple areas of contact model.

and for large values of N

$$A_2=\tfrac{1}{2}bhN^2=\tfrac{1}{2}Mbx^2=Bx^2. \tag{7}$$

Similarly, for the load and conductance,

$$P_2=\frac{ch^p}{(1+p)}N^{(1+p)}=\frac{Mc}{(1+p)}x^{(1+p)}=Cx^{(1+p)}, \tag{8}$$

$$G_2=\frac{dh^m}{(1+m)}N^{(1+m)}=\frac{Md}{(1+m)}x^{(1+m)}=Dx^{(1+m)}. \tag{9}$$

These results give the following generalized relationships between the contact area or conductance and the load:

$$\left.\begin{aligned} A_2&=B[P/C]^{[2/(1+p)]} \\ G_2&=D[P/C]^{[(1+m)/(1+p)]} \end{aligned}\right\}, \tag{10}$$

where $B=\frac{1}{2}Mb$, $C=[M/(1+p)]c$, and $D=[M/(1+m)]d$.

The detailed conclusions which result from Eqs. (6) and (10) are set out in Table I. This shows the theoretical dependence of the contact area A and conductance G on the load P for the single area and multiple areas cases.

4. Discussion

For a single area of contact, as shown in Sec. II.2, the conductance G is related to the load P by a relationship of the form

$$G = kP^n, \tag{11}$$

where k is a constant for the given set of conditions and n has values 1, $\frac{2}{3}$, $\frac{1}{2}$, or $\frac{1}{3}$, depending on the type of deformation and nature of contact resistance (Table I); these relationships have been observed experimentally. When other values of n are found, it has usually been assumed that this is due to a suitable combination of plastic and elastic deformation on the one hand and construction and film resistance on the other. However, when nominally mating surfaces are used this interpretation need not be valid, because the distribution of the total contact area between a number of individual areas introduces another factor which will affect the variation of conductance with load. Thus, it has been demonstrated above that other values of n in Eq. (11) can be deduced, for a given mechanism of deformation and a given type of contact resistance, by using a simple multiple-contacts model. The particular values of n which have been deduced are related to the model and, therefore, have no special significance; it can be shown that any model for which increasing load causes increases in the number and also in the size of the contact areas will give values of n higher than the corresponding single area case but not exceeding unity.

TABLE I. Theoretical relationships between contact area or conductance and the load.

Deformation	Type of conductance	Single area of contact	Multiple areas of contact
Elastic	...	$A \propto P^{2/3}$	$A \propto P^{9/8}$
Elastic	Film	$G \propto P^{2/3}$	$G \propto P^{9/8}$
Elastic	Constriction	$G \propto P^{1/3}$	$G \propto P^{9/8}$
Plastic	...	$A \propto P$	$A \propto P$
Plastic	Film	$G \propto P$	$G \propto P$
Plastic	Constriction	$G \propto P^{1/2}$	$G \propto P^{3/4}$

The few available measurements of the contact resistance of flats suggest that such larger values of n are found in practice. Thus, Holm (E.C.H. p. 84) gives results for clean metal plates of copper and nickel, the conductance/load curves (plotted on a log/log scale) having a slope of 0.6 which happens to be that predicted in Table I for constriction resistance elastic deformation conditions. His curves for clean crossed cylinders of the same metals at comparable and lighter loads have a slope of $\frac{1}{3}$ corresponding to the same conditions. The results for steel flats given by Bowden and Tabor, which are set out in Table II, give a value of n about 0.75.

From the experimental values of the conductance of flat surfaces, it is possible, assuming N equal circular contact areas each of radius "a" (and taking reasonable values of the flow pressure p_m, and the specific conductivity σ) to calculate the size and number of the contact areas involved.§ The results obtained from this type of analysis shown in Table II suggest that an increase in the load does increase both the number and the average size of these areas.

TABLE II. Estimates of number N and average radius a of the individual areas of contact involved in contact of steel flats (B.T. p. 31).

Load (g)	a (10^{-4} cm)	N
2000	5	3
5000	6	5
20 000	9	9
100 000	12	22
500 000	21	35

The foregoing discussion has been based on an extremely crude model, the surface being represented by a model involving a single scale of structure. A more realistic model would involve at least two scales of structure. For example, a typical well-ground surface may contain asperities of the order of 1 micron high on a base of some tens or hundreds of microns; however, this longer wavelength structure (which is revealed by profilometers and similar instruments) may have superimposed on it smaller asperities, which can be revealed only by instruments of higher resolution such as the electron microscope. The *general areas* where two such flat surfaces touch will be determined by the longer wavelength structure, but the individual real areas of contact may be formed by the deformation of the smaller asperities (E.C.H. p. 75, Moore,[3] 1948).

For a group of closely spaced contact areas, the constriction resistance resulting from the individual contacts is negligible compared with the larger constriction resistance of the general contact area of the group. Therefore, for clean metals the contact resistance should be determined primarily by the general contact areas, which are a function of the longer wavelength surface structure. Thus, experiments with crossed cylinders, even when the surface finish is quite rough and it is obvious that a continuous contact is not formed, give values of the conductance which agree quite well with the theoretical values obtained by assuming perfectly smooth contact members.[4] Again, for contact resistance measurements with clean metal flats, the deduced values of "a" (the radius of the average contact area) are of the order of 10^{-2} cm (E.C.H. p. 95, and Table II). This also is consistent with the suggestion that such contact resistance measurements give information about the general areas of contact.

In the process of sliding between nominally mating surfaces the frictional forces arise only from those

§ The effect of using the model of Sec. II.3 of this paper is that the size distribution of the contact areas which it involves introduces constant correction factors in the deduced values of N and a. See also E.C.H. p. 93.

[3] A. J. W. Moore, Proc. Roy. Soc. (London) **A195**, 231 (1948).

[4] M. Cocks, unpublished work carried out in this Laboratory.

parts of the surfaces which are within the range of intermolecular forces; this will involve the smaller contact areas. Therefore, as a guide to the interpretation of friction experiments, the evidence of contact resistance measurements should be treated with some reserve; for example, the deformation of the general contact regions might appear to be elastic on the basis of contact resistance measurements, whereas the friction between the same surfaces might be determined by the plastic deformation of the smaller superimposed asperities.

One may conclude that the limited evidence of contact resistance measurements is consistent with a model in which the number and average size of the contact areas increase with the load. Any more detailed discussion of the correct model to represent such surfaces must await further experimental evidence.

III. MECHANICAL WEAR

1. Introduction

In the following sections, the above calculation will be used as a starting point for a discussion of mechanical wear. Other authors have previously discussed this subject in terms of the real area of contact. Holm (E.C.H. p. 214.) assumed that the real area of contact is formed by the plastic deformation of contacting asperities and considered wear as an atomic process. On this basis, he has shown that the worn volume per unit sliding distance W is given by

$$W = ZP/p_m, \tag{12}$$

where Z is the number of atoms removed per atomic encounter and p_m is the flow pressure of the material. Assuming that the total real area of contact consists of N circular "a spots," each of radius a, and considering the worn volume as consisting of ξ-atomic layers removed from contacting "a spots," he shows that

$$Z = \frac{\xi\alpha}{2a}, \tag{13}$$

where α is the interatomic spacing.

Burwell and Strang[5] have discussed the existence of simple laws of wear (such as those forecast by Holm's theory) in the light of their experiments. They conclude that under their experimental conditions suitable combinations of materials can be found which show a simple wear behavior of the expected type, but that Holm's theory must be modified by replacing the concept of atoms removed at atomic encounters by the removal of wear particles at asperity encounters. Burwell and Strang join with Holm, and Rabinowicz and Tabor,[6] in suggesting that the deduction of simple laws of wear is dependent on the use of a model in which the average size of the individual contact area is constant, i.e., that increasing the load increases the number of contact areas and wear particles without affecting their average size. However, if this assumption is made, one would expect the wear to be proportional to the load whatever the mechanism postulated to produce the wear; conversely it follows that no theory embracing this assumption can be used to throw light on the mechanism of wear.

On the other hand, the present theory shows that the required relationships can be deduced without making any special assumptions as to the nature of the sliding surfaces. Although the model used in the earlier part of this paper is used as a starting point for the general theory, it is shown later that the more important conclusions are independent of the model used to represent the sliding surfaces.

In comparing the theory with the experimental evidence, two types of experiment have been considered. The first type concerns the transfer of material during the single passage of a slider over a flat surface. Rabinowicz and Tabor[6] have used a radioactive slider to measure this transfer and have shown that, for a number of metals, the amount of material transferred per unit sliding distance is proportional to the load. In the second, the more usual type of wear experiment, one is concerned with the material completely removed from both rubbing surfaces. In this type of experiment, it is well known that many mechanical and chemical factors may be involved; therefore, the simple theory of wear is not regarded as being applicable generally. However, the results of Burwell and Strang, where an inert lumbricant was used, show a simple behavior in which the wear is proportional to the load. Similar results have been obtained in some experiments not using a lubricant; these are briefly reported in Sec. III.3 below. In these experiments, and in those of Rabinowicz and Tabor, it seems reasonable to suggest that the wear mechanism may be particularly simple and suitable for theoretical treatment in terms of the real area of contact between the sliding surfaces.

2. A Simple Theory of Mechanical Wear

The assumptions which are made in deducing the theory are as follows:

(a) *Size and distribution of the total area of contact.* The multiple areas of contact distribution deduced in Sec. II.3. is used, but for a realistic model, asperities on both surfaces must be assumed.

(b) *Duration of any given contact.* The simplifying assumptions employed here are essentially those previously used by Blok[7] and Rabinowicz;[8] they also seem to be implicit in Holm's theory. At zero time, the location of the two areas forming a fully established

[5] J. T. Burwell and C. D. Strang, Proc. Roy. Soc. (London) **A212**, 470 (1952); J. Appl. Phys. **23**, 18 (1952).
[6] E. Rabinowicz and D. Tabor, Proc. Roy. Soc. (London) **A208**, 455 (1950).
[7] H. Blok, Inst. Mech. Engrs. (London) **2**, 222 (1937).
[8] E. Rabinowicz, J. Appl. Phys. **22**, 1373 (1951).

circular contact of radius a are shown in Fig. 3 at (a), while the position a short time later, after a sliding distance y, is shown at (b). After sliding a distance $2a$, the contact area under consideration has been reduced to zero [Fig. 3(c)]; in addition, it is assumed that at this moment a new similar contact area of radius a has just been fully established elsewhere in the surface. An alternative treatment involving the contact and deformation of two spherical asperities results in a similar value for the duration of the contact.

(c) *The shape of the worn particles.* Assuming that the contacts involved in rubbing the two surfaces are of the type indicated in (b), there are two simple assumptions which can be made regarding the volume δV of a given wear particle.

$$\text{(1) Layer removal,} \quad \delta V = \beta a^2, \tag{14}$$

where β is a constant. This implies that the depth of the material removed is constant; more specifically, it is independent of the load or the radius of the contact area.

$$\text{(2) Lump removal,} \quad \delta V = \gamma a^3, \tag{15}$$

where γ is a constant. This implies that the depth to which the material is torn is proportional to the radius of the contact area, i.e., statistically the shape of the wear particles is independent of their size.

(d) *The probability factor.* Not every contact results in a wear particle. It is assumed that a proportion K (independent of their size) does so. K may therefore be regarded as a probability factor typical of the combination of materials used, applying over a range of experimental conditions where the wear process remains of the same type.

If the wear rate i.e., the total wear per unit sliding distance for the whole surface, is W and the contribution from the areas under consideration is δW, then, from assumptions (b), (c), and (d),

$$\delta W = \mathrm{K}\delta V/2a,$$

and the wear rate W is given by

$$W = \Sigma \delta W,$$

where the summation is carried out for all the individual areas constituting the total area of contact at any time. Using a process similar to that previously employed in Sec. II.3, the wear rate W may be expressed as a function of the load P, the following expression being obtained:

$$W = \mathrm{K}E[P/C]^{[(1+q)/(1+p)]}. \tag{16}$$

C and p are as defined in II.3, and $E=[M/(1+q)]e$, where, for layer removal, $q=\frac{1}{2}$ and $e=\frac{1}{2}R^{\frac{1}{2}}\beta$ and for lump removal, $q=1$ and $e=R\gamma$. R is the radius of curvature of the asperities as before.

The theoretical relationships between the wear rate and the load for the four wear mechanisms are set out in Table III. Of these mechanisms the most likely is the removal of lumps from contact areas formed by plastic deformation. This is suggested by the results of friction experiments;[1] moreover, on the basis of the foregoing theory, it is this mechanism which gives a wear/load relationship in accord with the experimental evidence. Assuming hemispherical wear particles of the same radius as the contact areas, this mechanism gives the following expression for the wear rate:

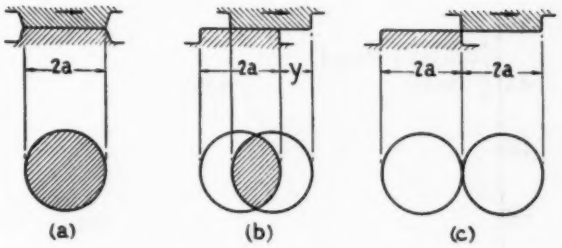


FIG. 3. Idealized representation of single contact in sliding surfaces. (a) Maximum contact area of radius a. (b) After sliding through distance y. (c) After sliding through a distance $2a$; contact area just reduced to zero.

$$W = \mathrm{K}P/3a. \tag{17}$$

This is similar to the equation of Holm [Eq. (12)] and is obtained essentially by replacing Holm's concept of removal of atoms by removal of wear particles.

It should be noted that the detailed "multiple areas of contact model," which has been assumed, is not an essential assumption in deducing Eq. (17); when the deformation is plastic, and wear occurs by the removal of lumps, it is only necessary to make assumptions, such as (b) and (c) above, about the duration of the contact and shape of the wear particle.

The essential argument may be summarized as follows: during sliding one cm, a number of contacts of different sizes will occur. Consider one such contact of radius a. We postulate: worn volume $\propto a^3$ and effective sliding distance $\propto a$; therefore, the contribution of this contact to the wear per cm of sliding $\propto a^2$; also load supported by contact $\propto a^2$. Therefore, for this contact, the contribution to the wear rate is proportional to the load supported by it. A similar argument applies to all other contacts, and the total wear rate is proportional to the load. For further discussion, see J. T. Burwell [Research (London) **6**, 25S (1952)] and J. F. Archard [Research (London) **6**, 33S (1952)]. Finally it should also be noted that the only effect of postulating

TABLE III. Theoretical relationships between the wear rate W and the load P.

Deformation	Particle shape	Relationship between W and P
Elastic	layer	$W \propto P^{2/3}$
	lump	$W \propto P^{4/3}$
Plastic	layer	$W \propto P^{3/4}$
	lump	$W \propto P$

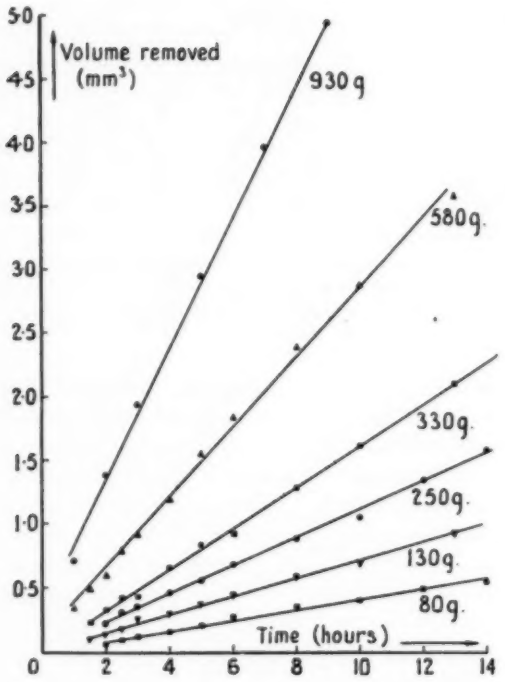


FIG. 4. Typical wear/time graphs for a stellite pin rubbing on a tool steel ring. These results represent a light type of wear with finely divided wear particles.

a different shape or size of wear particle (provided the shape is statistically independent of size) or a different duration of contact (provided the duration is proportional to the radius of the contact) is to change Eq. (17) by a constant factor.

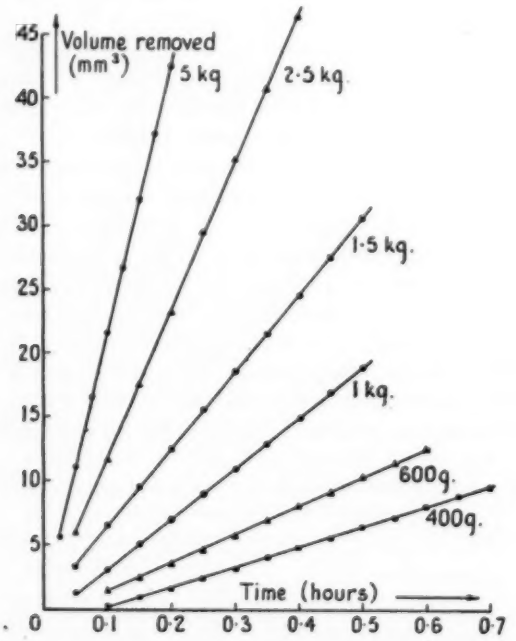


FIG. 5. Typical wear/time graphs for a 60/40 brass pin rubbing on a tool steel ring. These results represent a heavy type of wear with fairly large wear particles; a film of brass is established on the ring.

3. Discussion

Of the four possible mechanisms of wear set out in Table III, only that of lump removal from contact areas formed by plastic deformation represented by Eq. (17) will be considered in detail. The main conclusions which are derived from the theory for this case are:

(a) The wear rate is proportional to the load.
(b) The wear rate is independent of the apparent area of contact.
(c) Provided that K and p_m remain constant, the wear rate is independent of the speed of sliding.
(d) The theoretical value of the wear rate is independent of the model used to represent the surfaces.

There appears to be no evidence available for the adequate discussion of (c), but the experimental evidence for (a) and (b) will be considered.

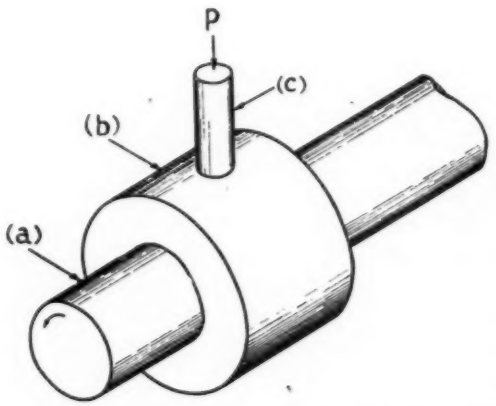


FIG. 6. Schematic diagram of pin and ring wear machine. (a) Shaft rotating at approximately 1500 rpm. (b) Ring 2.38 cm diameter. (c) Pin, 0.635-cm diameter, pressed under load P against ring.

Figures 4 and 5 show some results obtained in air with a simple pin and ring machine, the essential details of which are shown in Fig. 6. These wear/time graphs are linear, and it should be borne in mind that in this apparatus the apparent area of contact is not constant, but increases with the total wear. From the slopes of the wear/time graphs for these combinations, the wear rates, expressed as wear per unit sliding distance, have been obtained and are plotted against the load on a log/log scale in Fig. 7. The slopes of these wear rate/load graphs are 1.00 for brass and 0.98 for stellite (standard error 0.015 in each case). Thus, these experimental results (together with those of Burwell and Strang) suggest that simple laws such as (a) and (b) above can be found in practice. Similar results have also been found for other combinations of materials, including nonmetals, but in yet other cases simple laws of wear are not found or hold only over a limited range of conditions.

Taking the results shown in Fig. 7, experimental values of the wear rates are shown in Table IV; the theoretical expressions for the wear rates obtained from Eq. (17) are also given, together with the deduced values of the constant K. These two sets of results represent a heavy and light type of wear process which have been termed, respectively, "seizure wear" and "abrasive wear" by Mailander and Dies.[9] In the heavier type of wear (brass), Table IV suggests that about 0.2 percent of the contact areas contribute wear particles, whereas in the lighter type of wear (stellite), the figure is of the order of 0.02 percent. The results of Burwell and Strang[5] suggest that in the presence of an indifferent lubricant, the wear may be reduced by a further factor of 10^{-2} or 10^{-3}.

The same principles which have been applied above to wear may be used to discuss transfer of material during a single run over a surface. In Table V, the experimental values of metallic transfer given by Rabinowicz and Tabor[6] are compared with the theoretical expressions calculated from their data using Eq. (17). The following comments may be made.

(a) For similar metals, if every junction resulted in transfer, a value of K of 50 percent might be expected, since half the junctions would result in transfer in the opposite direction to that being investigated.

(b) In the results for similar metals, the size of the individual particles given by the authors suggests that they are of the same order of size as the expected value of the track width as calculated from the load and the flow pressure. Therefore, the values of K deduced in Table V probably indicate the degree to which the particles are spread out along the length of the track, each particle occupying a large proportion of the track width.

(c) With these provisos, the results suggest that for all cases of similar metals, except zinc, less than 10 percent of the junctions result in transfer; i.e., for over 90 percent of the track, where the load between the surfaces must be supported, no transfer takes place. In the case of dissimilar metals the deduced values of K are about 0.1 percent or less, but the experimental evidence is that this reduction is largely due to the reduced size rather than the number of the particles.

The interpretation of the deduced values of K, as indicating the proportion of contacts resulting in transfer (each particle being assumed hemispherical of the same radius as the contact area concerned), must be accepted with some reserve. As is suggested by the transfer experiments with dissimilar metals, the number of particles may be higher than is indicated by the value of K, but their size smaller. Nevertheless, the values of K indicate the degree to which the transfer or wear is reduced below the maximum expected value and so provide a qualitative basis for the assessment of the experimental results.

The position with regard to the various simple

[9] R. Mailander and K. Dies, Arch Eisenhüttenw. **10**, 203 (1943).

TABLE IV. Experimental and theoretical wear rates at 1-kg load.

Pin material	P_m kg/cm²	Wear rate cc/cm Theoretical [from Eq. (17)]	Experimental	Deduced value of K, percent
Brass	9.5×10^{3}	$3.5K\times10^{-5}$	6.2×10^{-8}	0.18
Stellite	6.9×10^{4}	$4.8K\times10^{-6}$	0.77×10^{-9}	0.016

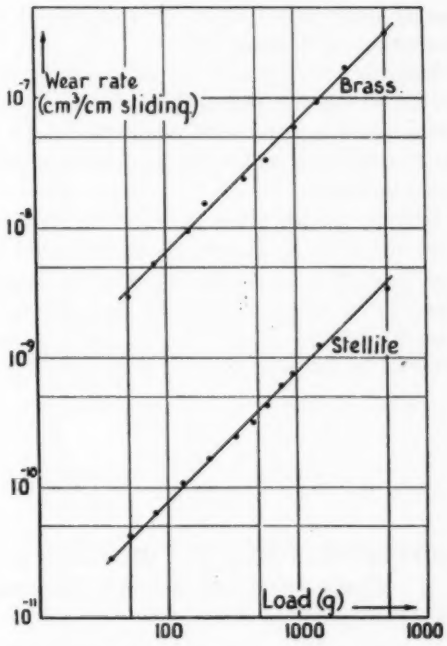


FIG. 7. Wear rate/load graphs for brass and stellite pins rubbing on tool steel rings.

TABLE V. Experimental and theoretical rates of transfer at 2-kg load.

(a) Similar metals.

Metal	P_m kg/cm²	Density g/cc	Transfer rates μg/cm Theoretical [from Eq. (17)]	Experimental (R and T)	Deduced value of K, percent
Cadmium	2×10^{3}	8.65	$2.9\ K\times10^{3}$	50	1.7
Zinc	3.8×10^{3}	7.14	$1.25K\times10^{3}$	200	16.0
Silver	4.3×10^{3}	10.5	$1.62K\times10^{3}$	20	1.2
Copper	9.5×10^{3}	8.95	$0.63K\times10^{3}$	20	3.2
Platinum	13.8×10^{3}	21.4	$1.03K\times10^{3}$	40	3.9
Mild steel	15.8×10^{3}	7.8	$0.33K\times10^{3}$	15	4.5
Stainless steel	21.7×10^{3}	7.8	$0.24K\times10^{3}$	5	2.1

(b) Dissimilar metals [values of p_m and density as in (a)].

Combination	Transfer rates μg/cm Theoretical [from Eq. (17)]	Experimental (R and T)	Deduced value of K, percent
Cadmium on mild steel	$2.9\ K\times10^{3}$	0.3	0.01
Copper on mild steel	$0.63K\times10^{3}$	1.0	0.15
Platinum on mild steel	$1.03K\times10^{3}$	1.5	0.15
Mild steel on copper	$0.55K\times10^{3}$	0.3	0.05
Platinum on silver	$3.3\ K\times10^{3}$	0.4	0.01

theories of wear which have been discussed may now be summarized as follows:

(1) Holm's first assumption of wear as an atomic process involving the removal of Z atoms per atomic encounter gives a wear rate proportional to the load but must be rejected on the basis of all the experimental evidence available.

(2) Holm's modification of this theory envisaging the worn material as atomic layers removed from the areas of contact is at variance with the experimental evidence, which suggests instead that the material is removed in lumps. This theory gives a wear rate proportional to the load, only if it is assumed that the average size of the contact areas and wear particles is constant.

(3) Burwell and Strang's general picture of wear as resulting from the removal of lumps at the contact of asperities is in much better accord with the experimental evidence, but the more detailed theory advanced in this paper shows that it is not necessary to assume that the average size of the contact areas or wear particles is constant.

The author wishes to thank Dr. W. Hirst for his continued interest and advice and his colleagues for many. helpful discussions. The author also thanks Dr. T. E Allibone, F.R.S., for permission to publish this paper.

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 8 AUGUST, 1953

Measurement of Elastic Constants at Low Temperatures by Means of Ultrasonic Waves—Data for Silicon and Germanium Single Crystals, and for Fused Silica

H. J. McSkimin
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received January 13, 1953)

Ultrasonic waves (shear or longitudinal) in the 10–30 mc range are transmitted down a fused silica rod, through a polystyrene or silicone one-quarter wavelength seal, and into the solid specimen. Measurement of reflections within the specimen yields values for velocities of propagation and elastic constants.

Data obtained over a temperature range of 78° to 300°K for silicon and germanium single crystals, and 1.6° to 300°K for fused silica are listed. For the latter, a high loss is noted, with an indicated maximum near 30°K.

1. INTRODUCTION

THE advantages of using high-frequency ultrasonic waves for measuring the elastic constants of solids—and in particular single crystals—have been pointed out by numerous investigators.[1–4] In the present paper application of principles outlined in reference 5 will be made to low-temperature measurements in the range −196°C to +25°C. Results for germanium and silicon single crystals will be listed, also data for fused silica. For the latter, a few attenuation measurements made with the use of liquid helium will be described also.

2. METHOD USED

2.1 Single Transducer Technique

Ultrasonic waves in the frequency range 10–30 mc can be transmitted readily down a fused silica rod as shown by Fig. 1. Either longitudinal or shear waves may be used.

The specimen is attached to the lower end of the rod with a plastic seal approximately $\frac{1}{4}$ wavelength in thickness. Details of this seal will be described later.

Only the specimen end need be cooled to the desired temperature; so that at all times the quartz crystal transducer is maintained near room temperature. Failure of the solder bond seal between crystal and rod, with resulting impairment of pulse transmission, is therefore avoided.

Waves entering the specimen through the seal are reflected back and forth, giving rise to a series of transmitted pulses T (Fig. 1), decaying exponentially with time. Because of the impedance transformation properties of the seal, only a small reduction in amplitude occurs for each reflection. Also, near the frequency for which the seal thickness is $\frac{1}{4}$ wavelength, the reflection phase shift error will be small.[5]

2.2 Determining the Number of Wavelengths in the Path-Velocity Calculation

At discrete frequencies, the emerging wave trains will be in phase, so that an observed pattern (envelopes only) occurs as shown by Fig. 2(a). The number of wavelengths in the complete path (for twice the specimen thickness) is given without ambiguity by

$$n_0 = f_0/\Delta f, \tag{2.1}$$

where f_0 is the critical frequency of interest, and Δf is

[1] H. B. Huntington, Phys. Rev. **72**, 321 (1947).
[2] J. K. Galt, Phys. Rev. **73**, 1460–1462 (1948).
[3] D. L. Arenberg, J. Acoust. Soc. Am. **20**, 1 (1948).
[4] W. C. Overton, Jr., and R. T. Swim, Phys. Rev. **84**, 758 (1951).
[5] H. J. McSkimin, J. Acoust. Soc. Am. **22**, 413–418 (1950).

the frequency separation between two consecutive critical frequencies. Accuracy in determining Δf is obtained by determining critical frequencies over a wide range of frequencies.

The velocity of propagation is then given by

$$V_0=\frac{2tf_0}{n_0-(\varphi_0/2\pi)}, \tag{2.2}$$

where t is the specimen thickness and φ_0 is the phase shift per reflection. In general, this phase shift term may be neglected at frequencies near the seal $\lambda/4$ frequency, so that the velocity of propagation is simply

$$V_0=2tf_0/n_0. \tag{2.3}$$

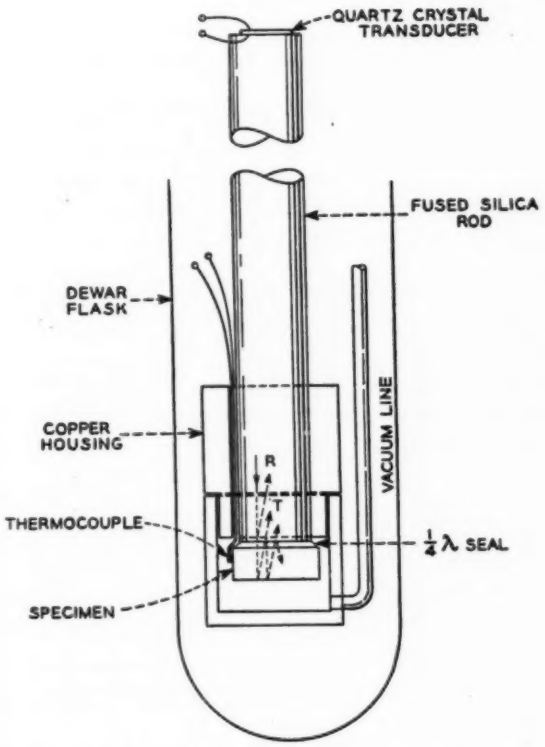


FIG. 1. Apparatus for measuring elastic constants.

Another useful pattern occurs for odd numbers of half-wavelengths as shown by Fig. 2(b).

3. DETAILS OF APPARATUS USED

3.1 Longitudinal Wave Unit

For longitudinal waves the fused silica rod of Fig. 1 was approximately 7.5 in. long by $\frac{7}{8}$ in. diam.[6,7] The ends were carefully aligned parallel and surface finished with No. 600 Carborundum.

[6] Material made by Nieder Company and supplied by Amersil Company, Inc., Hillside, New Jersey. Rods were clear and essentially strain free. Refer also to reference 7.

[7] E. S. Pennell, "Vitreous silica for ultrasonic delay line applications," to be published in Proc. Natl. Electronics Conf. Vol. 8, 1952.

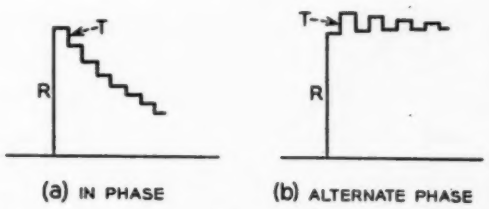


FIG. 2. Oscilloscope patterns for critical frequencies (envelopes).

A 25 mc x-cut quartz crystal was used as the transducer, being fastened to the rod by means of a silver paste-solder technique previously described.[5] Accurate measurements with this unit were limited to frequencies above 20 mc because of spurious "trailing" pulses generated at the rod surface. Additional roughening might have been resorted to for operation at lower frequencies.

3.2 Seals for Longitudinal Waves

For the temperature range between +25°C and −78°C, the most useful seal was found to be a 0.5-mil film of polystyrene with very viscous poly-α-methylstyrene[8,9] used as the binder. A 70°C assembly temperature was used with a pressure of approximately 25 lb/sq in.

Figure 3 shows measured values of R/T, the ratio of reflected to transmitted components (see Fig. 1) with an indicated $\lambda/4$ resonance at about 20 mc.[10,11] As the temperature is lowered, the liquid solidifies, and the polystyrene becomes stiffer so that the $\lambda/4$ point increases to about 23 mc at −78°C.

For the temperature range −78°C to −195°C, it was found that a suitable seal could be made with Dow Corning 200 Fluid of 200 000 cs viscosity (25°C). To provide control of the seal thickness, two tungsten wires 0.001 in. diam were stretched across the end of the fused silica rod. With a moderate pressure, the specimen (with excess of seal liquid and at room temperature) was pressed down until it touched the shims.

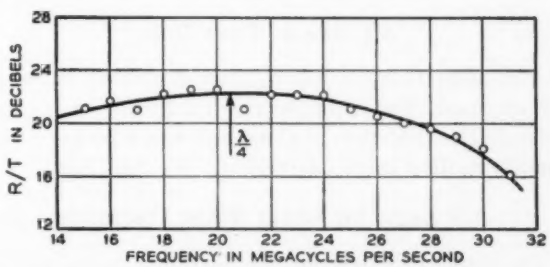


FIG. 3. Ratio of reflected to transmitted components for longitudinal waves along the 110 direction of a silicon crystal. (See text for seal.)

[8] Dow resin 276-V9, Blend 288. For properties, refer to reference 9.

[9] H. J. McSkimin, J. Acoust. Soc. Am. 23, 429–434 (1951).

[10] A description of measurements of this type is to be found in reference 11.

[11] Anderson Labs. Inc., "Interim report on solid supersonic delay lines," Dec. 20, 1950 to March 30, 1951.

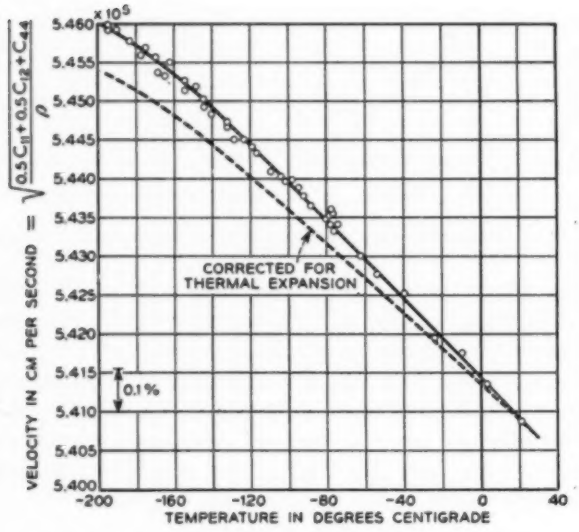


FIG. 4. Velocity of longitudinal waves in germanium single crystal—propagation parallel to 110 direction.

These wires were left in place, and the assembled unit (still in a vertical position) was cooled with dry ice until the seal was rigid enough to allow inverting the unit into a standard Dewar flask for further cooling by liquid nitrogen.

The quarter-wave frequency varied between 18 and 24 mc for the 1 mil shims depending on the temperature. Direct measurement of this frequency should be made from a comparison of R and T (Fig. 1) as previously noted.

3.3 Water Seal for Longitudinal Waves

An excellent low loss $\lambda/4$ seal can be obtained by the shim technique described above, using water or other suitable low viscosity liquid. A light pressure should be maintained on the specimen, and care should be exercised to keep liquid off the free face. Use of this technique in making loss measurements will be described later.

3.4 Shear Wave Unit

A construction similar to that described in Section 3.1 was used. The transducer was a 13-mc y-cut quartz crystal. The direction of the x-axis must be known, for particle motion takes place along this direction.

3.5 Seals for Shear Wave Operation

The same seals described in Section 3.2 were used for shear waves as well as longitudinal. The $\lambda/4$ frequencies were lower, in the range 10–15 mc.

3.6 Chamber for Specimen

In order to prevent moisture from condensing on the specimen, a vacuum-tight chamber was made of copper as shown by Fig. 1. With the cover removed, the end of the fused silica rod was completely accessible. Either a vacuum could be maintained, or the chamber could be filled with dry nitrogen.

Temperature equilibrium was obtained rather quickly because of the good heat conduction of the copper.

3.7 Temperature Measurement and Control

A small thermocouple of constantan and copper was placed on the specimen. Voltages were determined by means of a Leeds and Northrup Company No. 8657-C Double Range Potentiometer Indicator with uncertainties of about $\pm 1°C$.

In practice, the unit was cooled with liquid nitrogen to approximately $-195°C$ and then allowed to warm gradually. Measurements were made over the three- or four-hour period required for the temperature to rise to $-75°C$.

For temperatures between $25°C$ and $-75°C$, the unit was immersed in a dry ice-methanol mixture.

3.8 R.F. Oscillator and Gated Amplifier

The driving source for the transducer was a Measurements Corporation Model 80 Standard Signal Generator. The output was gated by means of a two stage amplifier, and provided pulse durations of from 1–50 microseconds.

Great care was taken to suppress the carrier in the output to minimize undesirable interference effects. Copper shielding was used throughout.

Frequencies were measured to ± 1 kc so that, in general, errors no greater than one part in 10 000 could be attributed to this source.

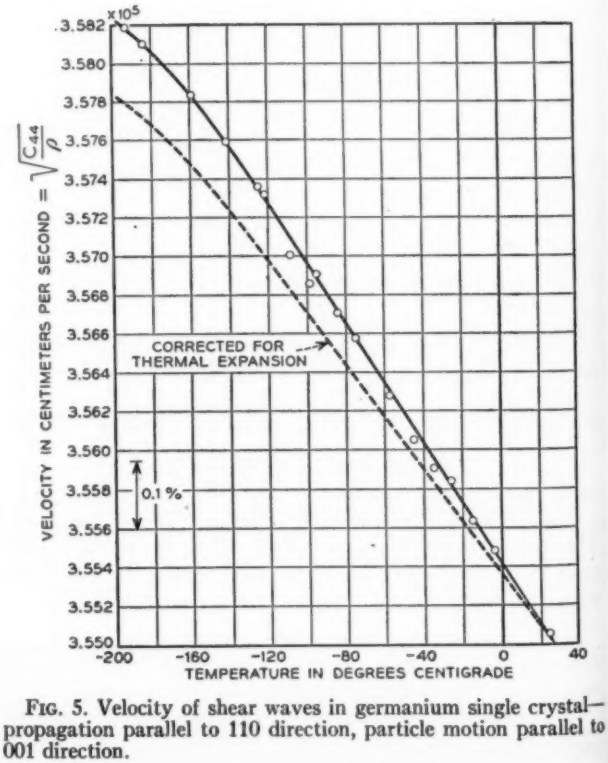


FIG. 5. Velocity of shear waves in germanium single crystal—propagation parallel to 110 direction, particle motion parallel to 001 direction.

3.9 Detection Equipment

A wide-band (4 mc) converter-amplifier was used with an IF frequency of 10 mc. After rectification, the signals were displayed on a DuMont Type 241 Oscilloscope.

4. RESULTS

4.1 Germanium Single Crystal

One of the germanium single crystals previously described[12,13] was measured for velocity of propagation, with results as shown by Figs. 4, 5, and 6. The data shown have been corrected for thickness-temperature changes in accordance with Fig. 7.[14]

Dimensions of the crystal were 1.9×1.3×0.6293 cm. The edge dimensions were left rough in order to scatter energy reaching the outer boundaries. The thickness

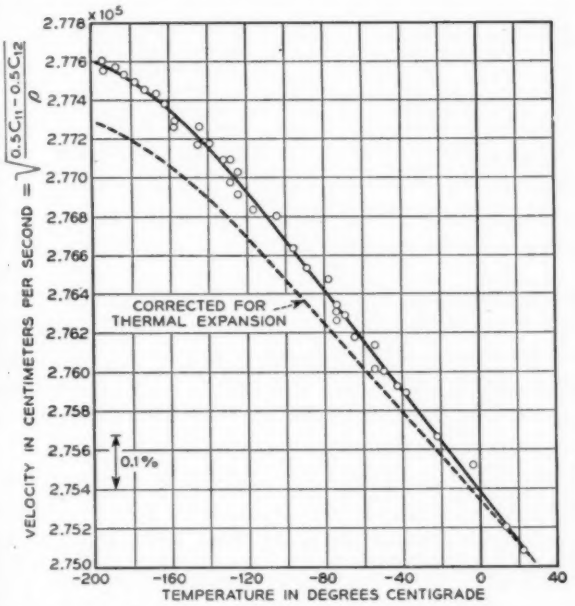


FIG. 6. Velocity of shear waves in germanium single crystal—propagation parallel to 110 direction, particle motion parallel to 110 direction.

(along a [110] direction) was not corrected for the slight pit penetration caused by the surface grind.[15] Direct measurement made by polishing the surface until pits disappeared indicated that errors in thickness dimensions were likely to be less than 0.1 percent. Overall uncertainties in velocity are estimated to be less than 0.5 percent.

A gradual bending over of the velocity curves at the lower temperatures is in evidence. This is consistent with a limiting value of zero for the slope at absolute zero.[2]

[12] Bond, Mason, McSkimin, Olsen, and Teal, Phys. Rev. **78**, 176 (1950).

[13] Teal, Sparks, and Buehler, Proc. Inst. Radio. Engrs. **40**, 906 (1952).

[14] Expansion data obtained by D. MacNair of Bell Telephone Laboratories.

[15] Roughening the surface by wet grinding with No. 600 Carborundum produced much stronger seals than a polished surface.

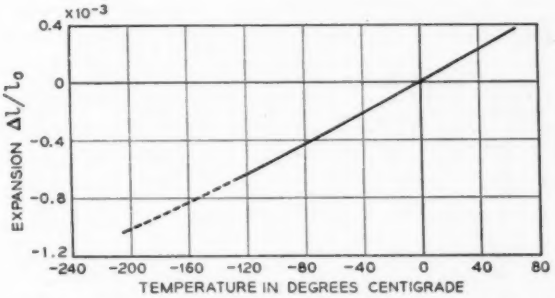


FIG. 7. Thermal expansion of germanium (D. MacNair).

A comparison of velocity data obtained in the present study with that for other germanium crystals measured is shown by Table I. Agreement in general is quite good.

4.2 Elastic Constants for Germanium Single Crystal

The three independent elastic constants for germanium were calculated from the velocities of propagation given by Figs. 4, 5, and 6. Results are shown by Figs. 8, 9, and 10. For these calculations, a density of 5.323 at 25°C was used.[16]

Measurements on single crystals of germanium have also been made by Mr. M. E. Fine with a low frequency (about 50 kc) resonance technique being used.[17,18] His

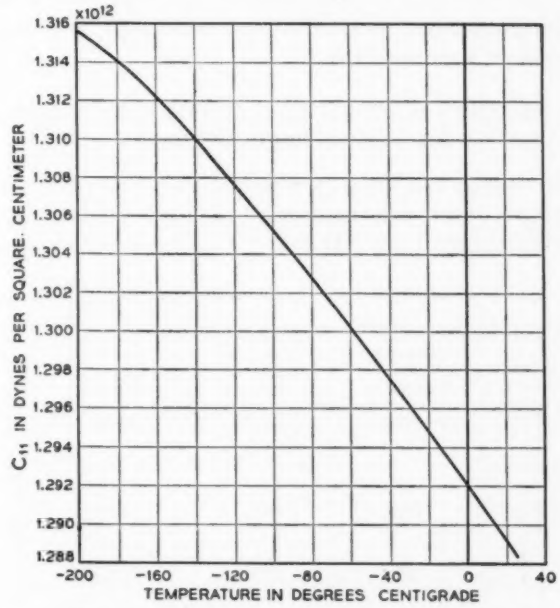


FIG. 8. C_{11} *versus* temperature for germanium.

[16] This value was supplied by Eagle Picher Lead Company. A value of 5.329 was determined with x-rays by W. L. Bond, Bell Telephone Laboratories. The 5.35 value used previously (reference 12) is apparently in error.

[17] M. E. Fine, "Dynamic methods for determining elastic constants and their temperature variation in metals," American Society for Testing Materials, Symposium on Determination of Elastic Constants, June, 1952.

[18] M. E. Fine, "Elasticity and thermal expansion of germanium between −195 and +275°C", submitted for publication in J. Appl. Phys.

TABLE I. Data for germanium single crystals.

Crystal designation	Direction of propagation	Direction of particle motion	Type wave	Temp. (°C)	Approx. freq. (mc)	Velocity (cm/sec)	Notes
B	110	110	Long.	26.4	10	5.410×10^5	Reported in reference 12
B				27	27	5.403	Remeasured
B				25	20	5.408	Present study
A				27	24	5.39	Reported in reference 12
B	110	001	Shear	28.3	10	3.547	Reported in reference 12
B				25	10	3.550	Present study
A				Rm	11	3.54	Reported in reference 12
B	110	$1\bar{1}0$	Shear	26.3	10	2.751	Reported in reference 12
B				25	10	2.751	Present study
A				30	10	2.75	Reported in reference 12
A	100	100	Long.	27	24	4.92	Reported in reference 12
C				25.3	27	4.9148	Imperfections present, *N*-type
D				29	27	4.9123	*N*-type, 1 hr. 900°C in N_2
D				26	14	4.914	
D				27	18	4.9136	λ/4 water seal used
E				27	17	4.9179	low res. *N*-type
F				27	17	4.9168	high res. *N*
G				27	18	4.9150	low res. *P*
H				27	17	4.9178	high res. *N*
I				30	17	4.9144	low res. *P*. (heat treated 50 min at 900°C in N_2)
A	100	001	Shear	30	7.5	3.54	Reported in reference 12

results are compared with those of the present study on Figs. 11 through 13. Considering that different samples, frequency ranges, and techniques were used, the close agreement is noteworthy.

4.3 Velocities of Propagation and Elastic Constants for Silicon Single Crystals

One of the two crystals reported on previously[19] was measured again with the thought of obtaining data over a temperature range. Detailed measurements over a wide frequency range, however, showed a dispersion effect for longitudinal waves not appreciated before. This particular crystal was approximately of 9 mm square cross section with two accurately parallel edge faces. At the lower frequencies, reflections from these faces caused interference effects with the "apparent" velocity varying as shown by Fig. 14. The use of a larger crystal with irregular edge boundaries was indicated, expecially in view of the large beam spreading due to the high velocity.

TABLE II. Data for silicon single crystals $T=25$°C.

Crystal No.	Direction of propagation	Direction of particle motion	Type mode	Velocity in cm/sec	Velocity-temperature coefficient	Notes
1	110	$1\bar{1}0$	Shear	4.682×10^5	-18×10^{-6}	
2				4.682	−47	(a)
3				4.673	−18.5	
2	110	001	Shear	5.834		
3				5.842	−17.9	$T=29$°C
1	110	110	Long.	9.167	−30	values calc. (b)
2				9.152		$T=30$°C(b)
3				9.129	−27.4	(c)
1	001	110	Shear	5.843	−21	(a)
1	001	001	Long.	8.474	−31	(a), (b)
1				8.420		(c)

Notes: (a) Velocity-Temp. Coeff. corrected in accordance with $\alpha=-3\times10^{-6}$ at 25°C.
(b) Velocity not corrected for geometric dispersion.
(c) Corrected for geometric dispersion.

[19] McSkimin, Bond, Buehler, and Teal, Phys. Rev. **83**, 1080 (1951).

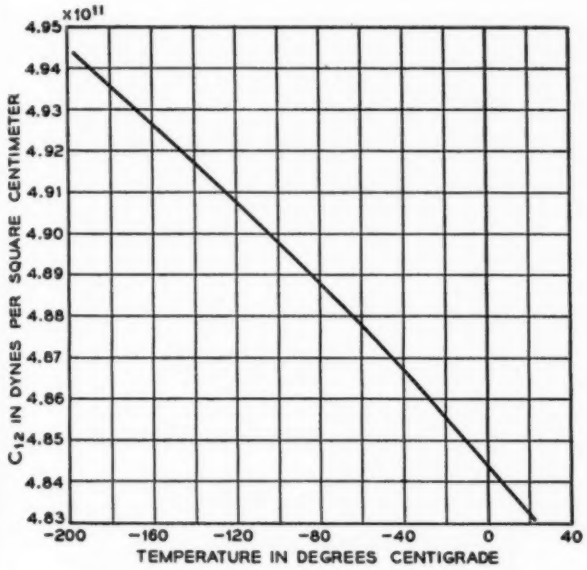


FIG. 9. C_{12} *versus* temperature for germanium.

Figure 15 shows similar data for a larger crystal with a thickness along the [110] direction of 0.9533 cm.[20,21] The cross dimensions were approximately 2.0×1.3 cm, with the edges quite irregular. It is seen that variations with frequency are quite small.

For shear waves, no significant variation with frequency was noticed.

Figures 16, 17, and 18 list data for velocities of propagation, again corrected for length expansion. The

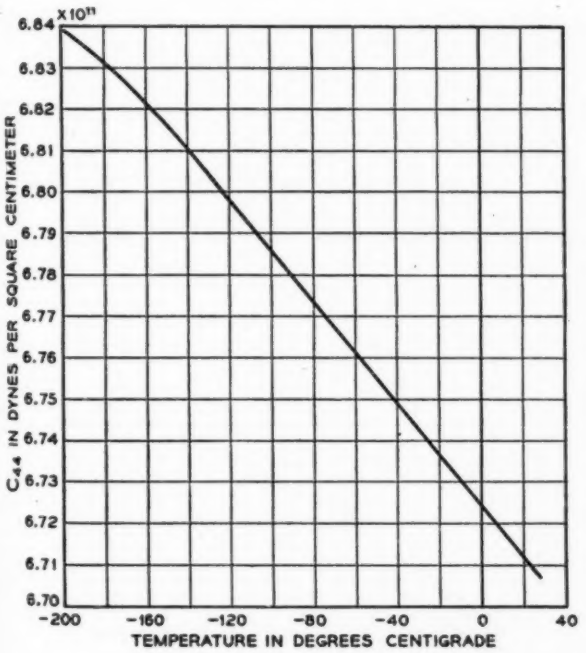


FIG. 10. C_{44} *versus* temperature for germanium.

[20] Crystal provided by E. Buehler and oriented by J. Andrus Refer to reference 17 for method of growing.

[21] G. K. Teal and E. Buehler, Bull. Am. Phys. Soc. 27, 14 (1952).

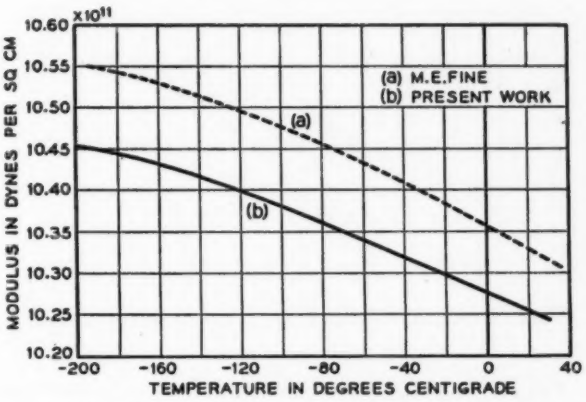


FIG. 11. Young's modulus for 100 direction in germanium.

data of Valentiner and Wallot was used.[22] (See Fig. 19 and reference 22.) The elastic constants are shown by Figs. 20, 21, and 22. A density of 2.331 at 25°C was used.

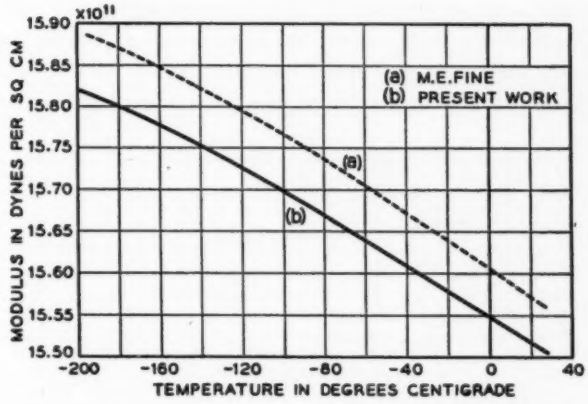


FIG. 12. Young's modulus for 111 direction in germanium.

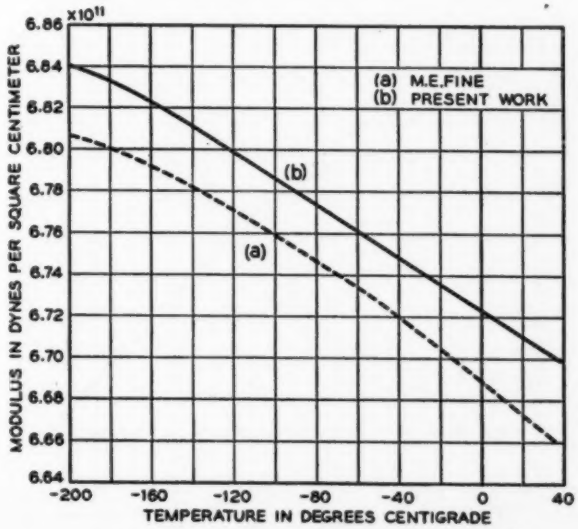


FIG. 13. Shear modulus C_{44} for germanium.

[22] S. Valentiner and J. Wallot, Ann. Physik 46, 837 (1915).

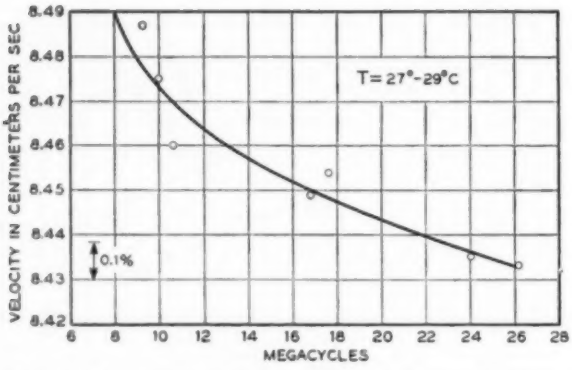


FIG. 14. "Apparent" velocity of longitudinal waves propagated along the 001 direction in silicon (cross dimensions approximately 9×9 mm, thickness 5.971 mm).

4.4 Attenuation Measurements

Because of the small specimen thicknesses used, it was not possible to obtain accurate determinations of attenuation for germanium and silicon over the temperature range of interest. Careful measurements at room temperature with direct evaluation of reflection loss using bayberry wax seals[5] indicated losses for the large silicon crystal tested to be less than 0.05 db per cm at

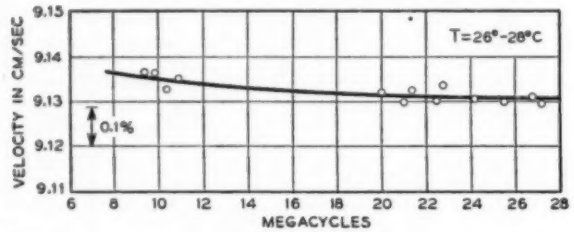


FIG. 15. "Apparent" velocity of longitudinal waves propagated along the 110 direction in silicon (cross dimensions approximately 13×20 mm and irregular, thickness 9.533 mm).

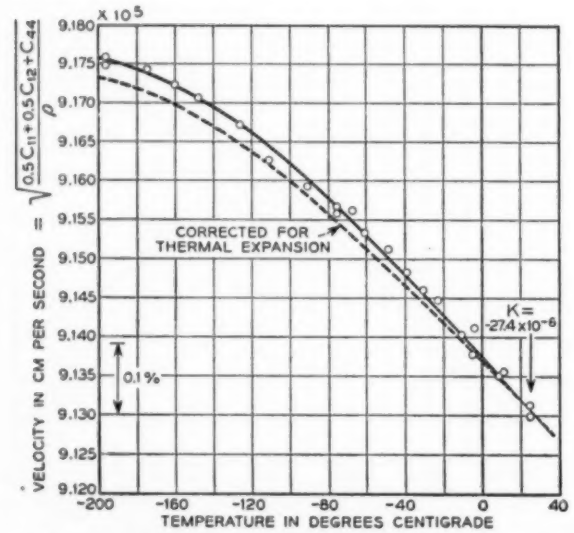


FIG. 16. Velocity of longitudinal waves in silicon single crystal—propagation parallel to 110 direction.

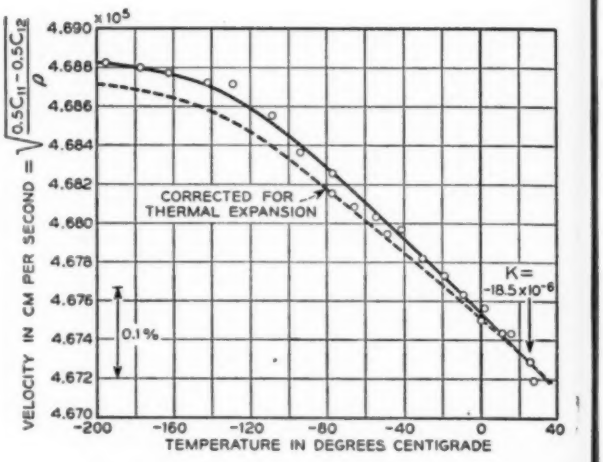


FIG. 17. Velocity of shear waves in silicon single crystal—propagation parallel to 110 direction, particle motion parallel to 110 direction.

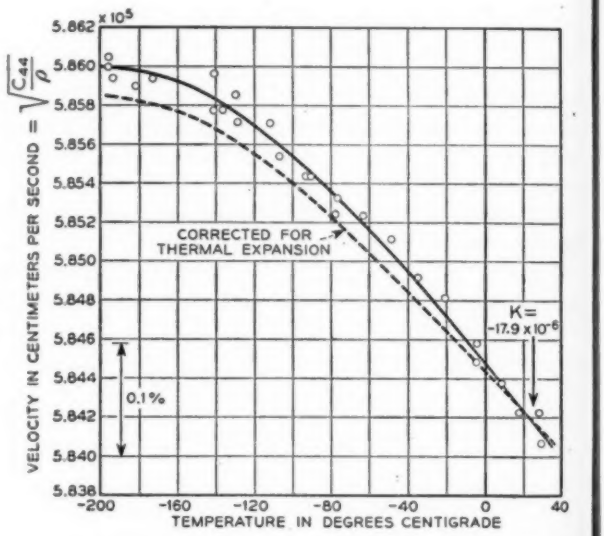


FIG. 18. Velocity of shear waves in silicon single crystal—propagation parallel to 110 direction, particle motion parallel to 001 direction.

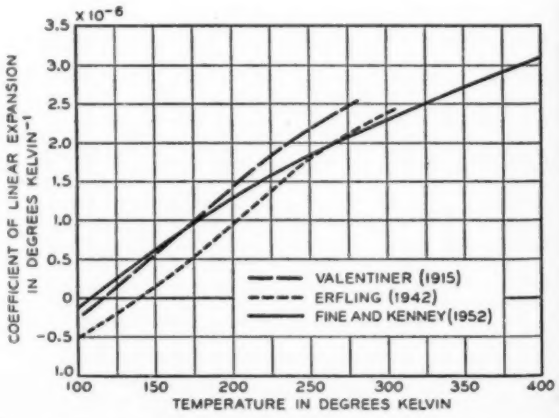


FIG. 19. Coefficient of linear expansion of silicon.

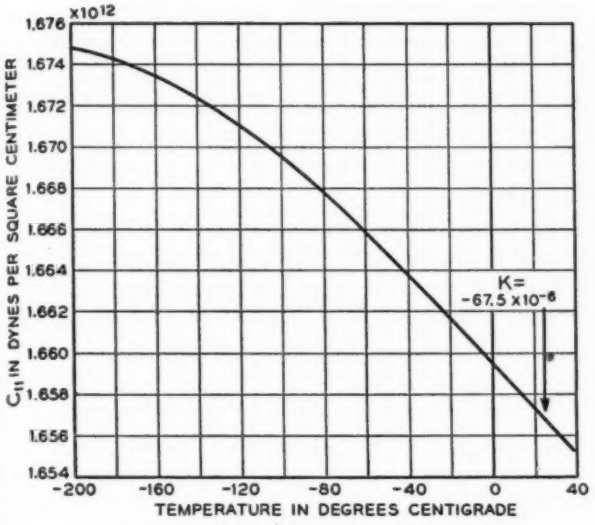


FIG. 20. C_{11} *versus* temperature for silicon.

frequencies as high as 15 mc for shear waves and 25 mc for longitudinal waves.

Room temperature measurements with longitudinal waves were also made for the germanium crystals (*E*, *F*, *G*, *H*, and *I*) of Table I. The water coupling technique described briefly in Sec. 3.3 was used, with total "round-trip" losses less than 0.2 db. Subtracting out a 0.11 db reflection loss due to the seal, the average loss for longitudinal waves along [100] (5 crystals) was approximately 0.06 db/cm at 18 mc.

The advantage of this method over other water coupling schemes[23] is apparent from the very low reflection loss at the seal. It would appear feasible to extend this method with the use of other liquids to cover a temperature range of at least −100°C to +100°C.

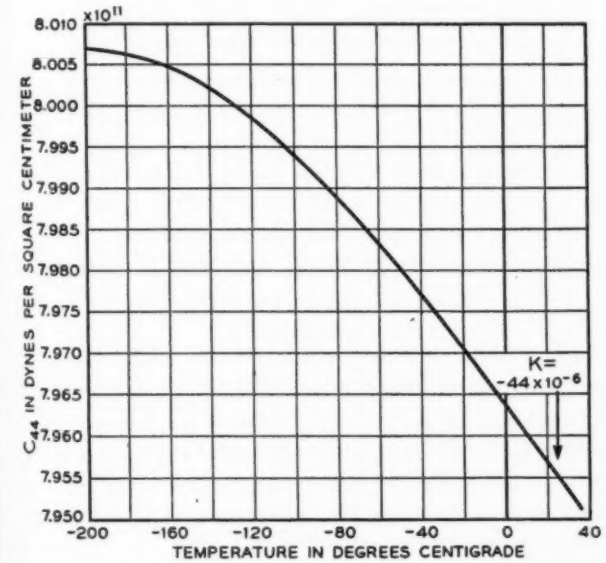


FIG. 21. C_{44} *versus* temperature for silicon.

[23] W. Roth, Quarterly Progress Report, M.I.T. Research Lab. of Electronics, Oct. 15, 1947.

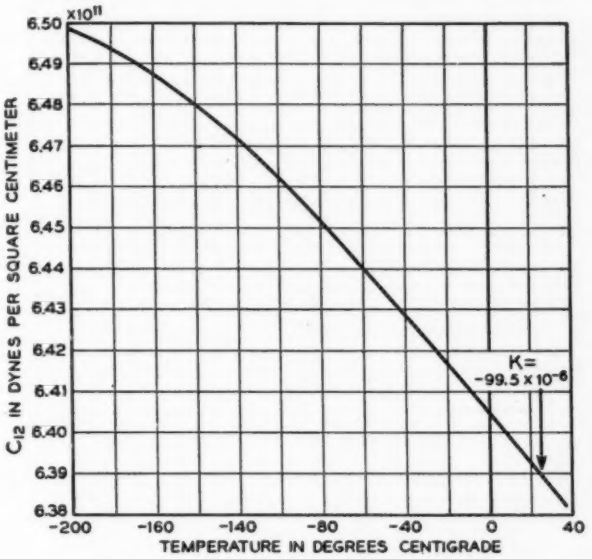


FIG. 22. C_{12} *versus* temperature for silicon.

4.5 Attenuation for Fused Silica

It was noticed that at temperatures below −150°C the response of the units fell off appreciably. With about 6 in. of the longitudinal wave rod[24] encased in a copper

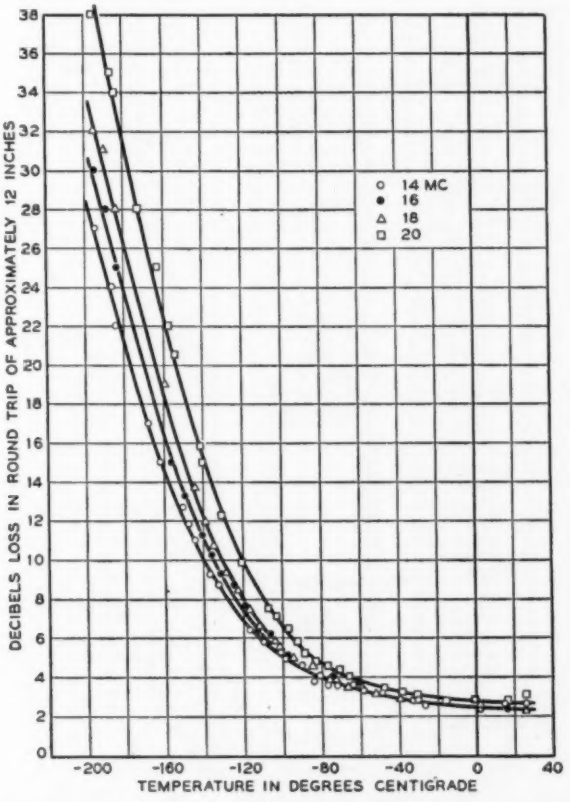


FIG. 23. Loss for longitudinal waves in fused silica rod. (Data include reflection loss at crystal end and beam spreading loss, estimated 2 db total.)

[24] The "effective" length covered was probably slightly greater than 6 in. because of heat conduction within the rod.

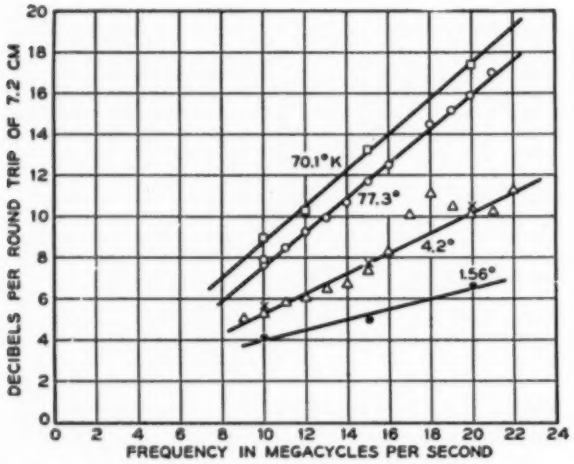


FIG. 24. Loss for shear waves in fused silica rod as a function of frequency and temperature. ($l = 3.60$ cm, diam $= 2.25$ cm).

sleeve and immersed in liquid nitrogen, the pulse decay rate was measured as a function of frequency with the results shown in Fig. 23. Similar data for shear waves was obtained.

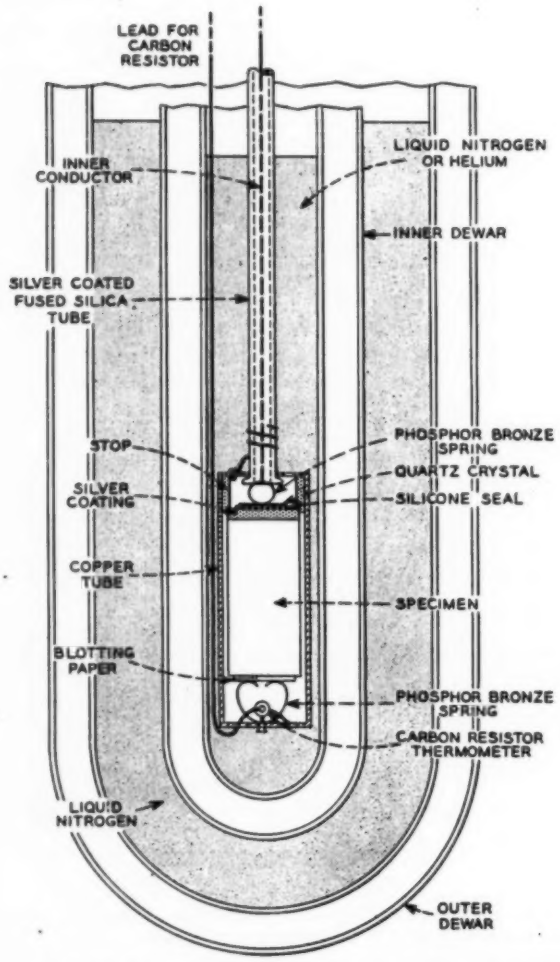


FIG. 25. Apparatus for measuring loss in fused silica at low temperatures.

Additional measurements were made with other samples of fused silica of similar quality (see reference 6). Figure 24 shows shear wave data for one of length 3.60 cm and diameter 2.25 cm. For these measurements the specimen was immersed in either liquid nitrogen or liquid helium, with the experimental arrangement of Fig. 25 being used.

The 18 mc y-cut quartz crystal transducer was attached to the specimen with a very viscous silicone,[25] a pressure of approximately 10 lb per sq in. being used. It was found advisable to roughen the silver coating on the end of the specimen with fine emery paper to increase the strength of the seal. Preparation of this silver coating was the same as described in reference 5.

The temperature was varied over narrow limits by varying the pressure over the liquid nitrogen or helium.

As seen from Fig. 24, the loss is approximately linear

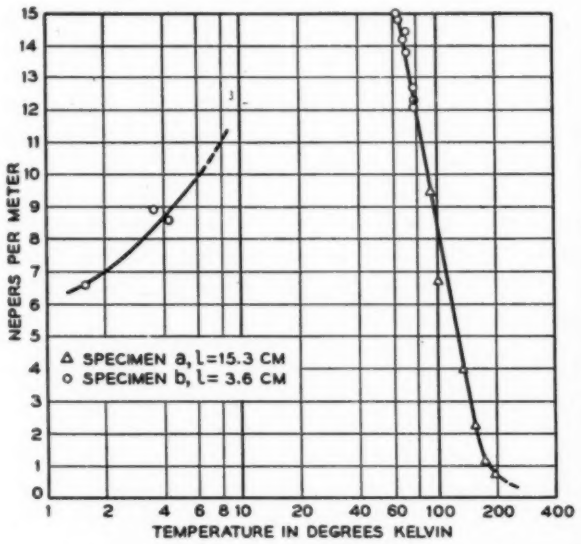


FIG. 26. Attenuation of shear waves in fused silica as a function of temperature (frequency $= 10$ mc).

with frequency over the two-to-one range used. The hump in the curve for 4.2°K was found to result from loss of energy to the electrical termination near the 18 mc crystal resonance frequency. Lowering the termination minimized this effect as can be seen from the 77.3° data.

Figure 26 shows a plot of shear wave attenuation as a function of temperature for 10 mc. A maximum is indicated near 31°K. From the shape of the loss curve, it would appear that a relaxation process exists. Correlation of data for the two different types of wave motion indicates that whatever this process is, it is primarily sensitive to shear distortion.[26]

[25] Dow Corning 200 Fluid, viscosity 25 000 poises.

[26] The mechanism of this relaxation process will be discussed in a forthcoming paper by O. L. Anderson and H. J. McSkimin.

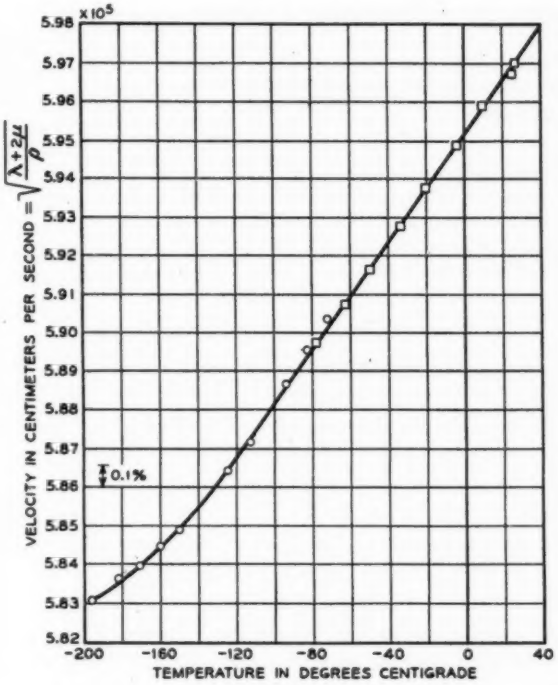


FIG. 27. Velocity of longitudinal waves in fused silica as a function of temperature.

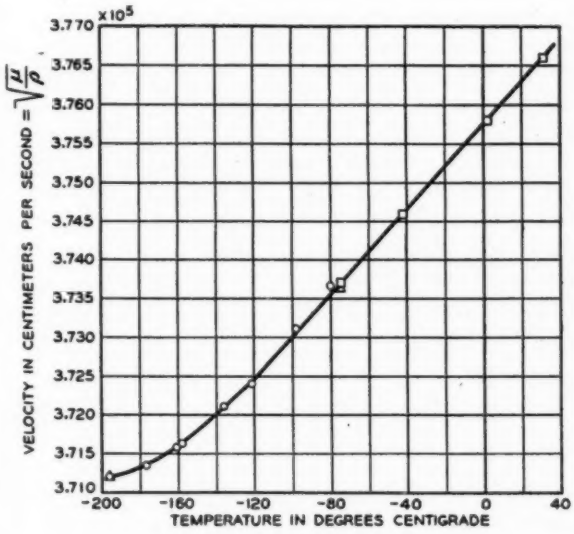


FIG. 28. Velocity of shear waves in fused silica as a function of temperature.

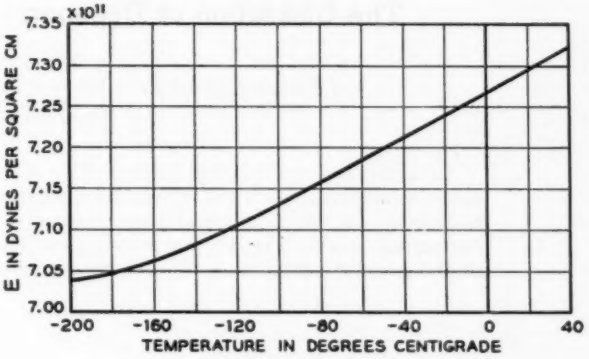


FIG. 29. Young's modulus E for fused silica (frequency = 20 mc, approx).

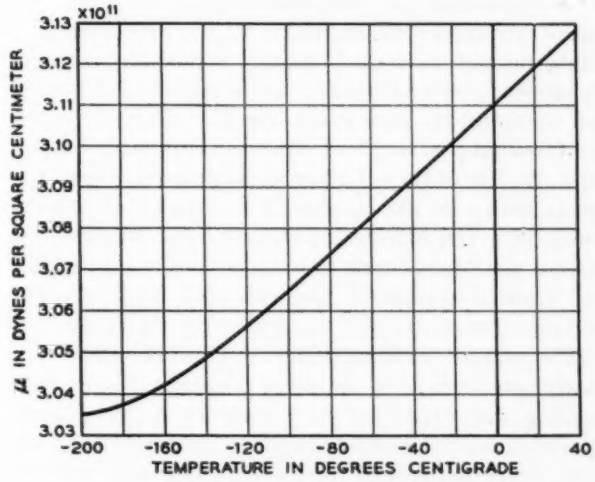


FIG. 30. Shear modulus μ for fused silica (frequency = 12 mc, approx).

4.6 Velocities of Propagation and Elastic Constants for Fused Silica

Figures 27 through 30 show the variation of velocity and elastic constants for fused silica as a function of temperature. A density of 2.203 g/cc was used.[27]

In addition to acknowledgments in the references, the author wishes to thank Mr. M. E. Fine for permission to reproduce his data for expansion coefficients and germanium elastic constants, and also to express appreciation for measurements and calculations made by Mr. T. B. Bateman.

[27] R. B. Sosman, *The Properties of Silica* (Chemical Catalog Company, Inc., New York, 1927) p. 361.

The Utilization of Degenerate Modes in a Spherical Cavity*

MALCOLM R. CURRIE
Electronics Research Laboratory, University of California, Berkeley, California
(Received January 21, 1953)

A general expression is given for calculation of the coupling coefficients between an arbitrary number of modes simultaneously excited in a cavity resonator. The mechanism of energy conversion between modes is discussed for several practical coupling devices. The technique is applied to the unlimited numbers of degeneracies available in spherical geometry and is shown to offer many interesting possibilities. An experimental model of a spherical cavity is described and shown to demonstrate the theory.

INTRODUCTION

MUCH interest has recently been expressed in the technique of coupling energy between the degenerate modes of oscillation which may exist in a cavity resonator. In particular, it is possible to couple the modes in such a manner that the resulting configuration simulates the prescribed behavior of an equivalent two-terminal-pair network over a limited frequency band. W. Lin of these Laboratories succeeded in coupling as many as five modes of a circular cylindrical cavity in a chain from input to output and thereby realized the filter characteristics otherwise obtainable only from a system of five cascaded cavities.[1] The structure was compact, light, and easily adjustable.

The purpose of this paper is twofold. First, the possible methods of establishing the desired intermode coupling are expanded and included in a general expression for the mutual coupling coefficient between an arbitrary number of modes simultaneously excited in a single cavity. The presentation unifies the results of all earlier work into a general theory and includes the treatment of a new device for obtaining coupling between modes. Second, the interesting case of a spherical cavity, with its unlimited numbers of degeneracies, is investigated for application of the general technique.

THE INTERMODE COUPLING TECHNIQUE

A cavity resonator having both input and output couplings can be conveniently regarded as a four-terminal network possessing poles at each of its natural modes of oscillation.[2] If appropriate values of inductance and capacitance are ascribed to each mode, the equivalent lumped element representation of the cavity takes the form of an infinite number of parallel-resonant circuits, each coupled directly to the input and output terminals. Any perturbations of the cavity geometry will destroy the independence of some of the modes and act to couple energy between them. This can be represented in the equivalent network by ideal mutual couplings with suitably defined coupling coefficients.

If, at a particular frequency ω_0, the cavity can support N degenerate modes, the presence of the couplings serves to slightly shift the frequencies of these modes relative to one another and thus removes their pre-existing degeneracy. Consider now the external behavior of the system around frequency ω_0. Only those modes having resonance frequencies in the vicinity of ω_0 need be retained in the equivalent circuit. Therefore, when other points of resonance are well separated from those of the perturbed degenerate modes, the network reduces to a configuration of N resonant circuits, mutually coupled to one another by varying amounts and with input and output terminals.

In this section an approximate expression for the coupling coefficient between degenerate modes is presented for several practical coupling devices. This permits a quantitative means for evaluating the equivalent circuit of such a system and conversely, allows such a system to be designed for prescribed external frequency behavior. The method constitutes a circuit approach to the field problem and is very useful in many complicated configurations.

The total fields in the cavity are taken as the sums over the N orthogonal modes of oscillation which are assumed to be excited,

$$\bar{E}=\sum_{i=1}^{N} e_i\bar{E}_i, \quad \bar{H}=\sum_{i=1}^{N} h_i\bar{H}_i. \tag{1}$$

$\bar{E}_i$ and $\bar{H}_i$ are eigenfunction solutions to the wave equation subject to boundary conditions imposed by the unperturbed cavity geometry and are normalized to unity over the volume of the cavity. The coefficients in the expansions are given by

$$e_i=\int_V \bar{E}\cdot\bar{E}_i dV, \quad h_i=\int_V \bar{H}\cdot\bar{H}_i dV. \tag{2}$$

Following a procedure employed by Slater,[3] Eq. (1) can be substituted into Maxwell's equations, yielding

* This work was done at the Electronics Research Laboratory of the University of California and was supported in part by the U. S. Office of Naval Research.

[1] W. Lin, J. Appl. Phys. **22**, 989–1003, (1951).

[2] S. A. Schelkunoff, Proc. Inst. Radio Engrs. **32**, 83 (1944).

[3] J. C. Slater, *Microwave Electronics* (D. Van Nostrand Company, Inc., New York, 1950), Chapter 4.

a series of oscillation equations for the coefficients h_i,

$$\epsilon\mu\frac{d^2}{dt^2}h_i+\beta_i^2h_i=\beta_i\int_{S'}(\bar{n}\times\bar{H})\cdot\bar{E}_i dS' + \epsilon\frac{d}{dt}\int_S(\bar{n}\times\bar{E})\cdot\bar{H}_i dS, \quad (3)$$

where the boundary conditions on $\bar{E}_i$, $\bar{H}_i$ are postulated as $\bar{n}\times\bar{E}_i=0$ on S, $\bar{n}\times\bar{H}_i=0$ on S', and $\bar{n}$ is a unit vector normal to the surface, projecting into the cavity. Also, $\beta_i=\omega_i(\mu\epsilon)^{\frac{1}{2}}$.

In the case of ideal cavity geometry with no losses, the right side of Eq. (3) vanishes and h_i is then determined as a function of time by a differential equation of simple harmonic motion. However, if small perturbations are introduced, the terms at the right in Eq. (3) will have finite values and will assume the role of an external driving force. The independence of the orthogonal modes will be destroyed and interaction between the modes may take place.

A perturbation procedure will be applied to Eq. (3) to allow for the presence of very small coupling devices placed within the cavity volume. Thus, a correction to Eq. (3) is obtained which accounts for the effects of the perturbations to first-order approximation. In analogy to coupled circuit theory, relationships for the coefficients of coupling between the existing modes can then be set up.

Three principal types of perturbations are considered. The first type consists of small volumes of perfect conductor placed at critical points within the cavity. For example, small metal slugs can be suspended inside the cavity by threads or indentures of the cavity wall can be obtained by inserting probes through the side-walls. The second type of coupling device consists of small, thin loops placed in positions of high magnetic field. As will be seen, this technique can establish a type of coupling not possible with probes alone. The above-mentioned methods are sufficient to establish any physically realizable intermode coupling. However, energy must also be coupled into and out of the cavity, and it is often convenient to accomplish this by the use of irises in the side-walls. Therefore, it is useful to approximately account for the presence of irises and their effect on the frequency of oscillation.

Referring to Eq. (3), only the surface integral over S' need be included for the probes and loops; both integrals must be included over the irises. In evaluating the integrals it is assumed that the perturbations to the fields are small so that, to a first-order approximation, the actual fields can be replaced by the original fields as given in Eq. (1). This procedure is, of course, not strictly correct, but experimental evidence suggests that it yields quite accurate results if the coupling devices are kept small.

The integrals over the surfaces of the probes can be put into convenient form by transformation to integrals over the perturbing volumes.[3] For the loops, the surface integrals are taken over tubular surfaces surrounding the wires of the loops, as shown in Fig. 1. In performing the integration, one first integrates around the circumference of the wire (using $d\bar{s}$) and then along the loop using length element $d\bar{l}$. The latter integral can be transformed to a surface integral over the area A enclosed by the loop itself. Details of this calculation follow very closely those of a calculation given by Hansen[4] for the excitation of a cavity by a loop. The results of Bethe's theory of diffraction by small holes[5] are employed in estimating the contributions to Eq. (3) due to the irises. The actual fields in the irises are approximated by the total normalized fields $\bar{E}_0$ and $\bar{H}_0$ evaluated at the holes.

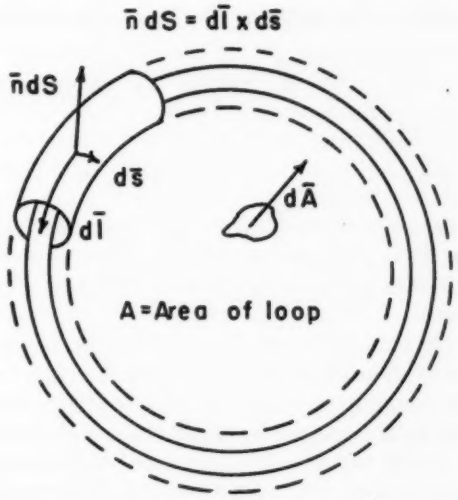


FIG. 1. Details of integration over the surface of a coupling loop.

Substituting the results of the above calculations into Eq. (3) and assuming that $\omega_i\cong\omega_0$ for the N "degenerate" modes, one obtains a set of N equations relating the relative amplitudes of the excited modes of the form

$$\left[\frac{\omega^2}{\omega_0^2(1+\Delta)}-1\right]h_i-\sum_{\substack{j=1\\ j\neq i}}^{N}h_jK_{ij}=0, \quad i=1,2,\cdots,N, \quad (4)$$

where

$$\Delta=\sum_{\text{probes}}\int_{V_P}(H_i^2-E_i^2)dV_P + \sum_{\text{loops}}\oint\bar{H}_i\cdot d\bar{s}\int_A\bar{H}_i\cdot d\bar{A} + \sum_{\text{irises}}2(PE_{i0}^2-MH_{i0}^2), \quad (5)$$

with $M=4/3(d/2)^3$ and $P=\frac{2}{3}(d/2)^3$, as defined by Bethe. A coupling coefficient K_{ij} between the ith and

[4] W. W. Hansen, J. Appl. Phys. 9, 654–663 (1938).
[5] H. A. Bethe, "Theory of Diffraction by Small Holes," M.I.T. Radiation Laboratory Report 128 (V–155), Jan. 23, 1942.

jth modes is thus given by the approximate relation

$$K_{ij}=\frac{\sum_{\text{probes}}\int_{V_P}(\bar{H}_i\cdot\bar{H}_j-\bar{E}_i\cdot\bar{E}_j)d+V_P\sum_{\text{loops}}\oint\bar{H}_j\cdot d\bar{s}\int_A\bar{H}_i\cdot d\bar{A}+2\sum_{\text{irises}}(P\bar{E}_{i0}\cdot\bar{E}_{j0}-M\bar{H}_{i0}\cdot\bar{H}_{j0})}{1+\Delta}. \quad (6)$$

Thus for an arbitrary array of very small coupling devices, it is possible to calculate the mutual effects of the modes upon one another in terms of a coupling coefficient.

The physical meaning of Eq. (6) should be examined in order to make clear the coupling mechanism. Consider first the coupling between the ith and jth modes which is set up by a single probe or piece of conductor placed in the cavity volume. It is apparent that if magnetic field components exist at the position of the probe and if the electric field components are negligible, a magnetic or inductive coupling effect will be established. Furthermore, Eq. (6) shows that maximum coupling will be realized if the magnetic field components are collinear at the probe position and that the coupling will vanish if the fields are at right angles. Similarly, a capacitive coupling effect is produced by the presence of nonperpendicular electric field components at the probe. In general, then, the fields of the two modes will experience a net interaction and hence a net coupling of energy from one to the other provided the fields are not perpendicular at the perturbation. The mechanism of energy conversion can be visualized in terms of driving surface currents set up by the discontinuity of the total magnetic field at the position of the probe, as indicated by the $(\bar{n}\times\bar{H})$ term in Eq. (3). This type of coupling was employed by Lin[1] in his studies and has been used in other microwave components for the purpose of coupling energy between electric fields.[6]

Equation (6) shows that a different type of coupling can be established by proper orientation of a thin loop. In particular, it is possible to set up interaction between two modes having perpendicular magnetic field components at the position of the loop. The maximum effect will occur when the plane of the loop is inclined at some angle to each of the magnetic field vectors of the modes to be coupled. Here the source of excitation of the second mode can be pictured in terms of a current induced in the loop by the varying flux linkage due to the first mode.

An important application of the foregoing results is the realization of a chain of successively coupled modes from input to output. The equivalent circuit of such a configuration is a cascade of coupled resonant circuits and thus represents a multistage band-pass filter. For this application, unique points in the cavity must be be determined at which one mode can be coupled to the next in the chain without introducing random couplings among any of the other modes or between any mode and the externally connected input and output lines.

[6] J. R. Pierce, Proc. Inst. Radio Engrs. **37**, 152–155 (1949).

DEGENERATE MODES IN A SPHERICAL CAVITY

The perfect geometrical symmetry of a spherical resonator allows its modes to possess a high degree of degeneracy, i.e., many independent field configurations having the same natural frequency of oscillation. Indeed, no other geometry offers greater possibilities. In this section the nature of spherical degeneracies will be discussed along with some examples of how they can be utilized according to the principles of the foregoing section to yield the frequency behavior of complex networks over a limited band. In addition, an experimental model of a spherical cavity excited in several modes is described in order to demonstrate the general technique and to show the practicability of constructing spherical resonators.

The solution to the field equations in spherical coordinates and the application of appropriate boundary conditions at the bounding surface $r=a$ determine the characteristic values β_{ln} and β_{ln}' for the TE and TM modes, respectively,

$$TE: \qquad J_{n+\frac{1}{2}}(\beta_{ln}a)=0, \quad (7)$$

$$TM: \quad \left\{\frac{d}{dr}[(\sqrt{r})J_{n+\frac{1}{2}}(\beta_{ln}'r)]\right\}_{r=a}=0, \quad (8)$$

where the integer index l denotes the lth root of (7) or (8). The wavelengths of natural resonance are given by

$$\lambda_{ln}=2\pi/\beta_{ln}, \quad \lambda_{ln}'=2\pi/\beta_{ln}'. \quad (9)$$

Explicit expressions for the field components of the TE_{lmn} mode, normalized to unity over the sphere's volume, are

$$E_\theta=A\left(\frac{a}{r}\right)^{\frac{1}{2}}\frac{1}{\sin\theta}P_n^m(\cos\theta)\frac{d}{d\phi}\begin{bmatrix}\cos\\ \sin\end{bmatrix}m\phi\Big]J_{n+\frac{1}{2}}(\beta_{ln}r),$$

$$E_\phi=A\left(\frac{a}{r}\right)^{\frac{1}{2}}\frac{d}{d\theta}P_n^m(\cos\theta)\left\{\begin{matrix}\cos\\ \sin\end{matrix}\right\}m\phi\Big\}J_{n+\frac{1}{2}}(\beta_{ln}r),$$

$$H_r=A\frac{n(n+1)}{\beta_{ln}a}\left(\frac{a}{r}\right)^{\frac{1}{2}}P_n^m(\cos\theta)\left\{\begin{matrix}\cos\\ \sin\end{matrix}\right\}m\phi\Big\}J_{n+\frac{1}{2}}(\beta_{ln}r),$$

$$H_\theta=A\frac{a^{\frac{1}{2}}}{\beta_{ln}r}\frac{d}{d\theta}P_n^m(\cos\theta)\times\left\{\begin{matrix}\cos\\ \sin\end{matrix}\right\}m\phi\Big\}\frac{d}{dr}[(\sqrt{r})J_{n+\frac{1}{2}}(\beta_{ln}r)], \quad (10)$$

$$H_\phi = A\frac{a^{\frac{1}{2}}}{\beta_{ln}r}\frac{1}{\sin\theta}P_n^m(\cos\theta) \times\frac{d}{d\phi}\begin{bmatrix}\cos\\ \sin\end{bmatrix}m\phi\Big]\frac{d}{dr}[(\sqrt{r})J_{n+\frac{1}{2}}(\beta_{ln}r)],$$

$$E_r=0,\quad A=\frac{1}{V^{\frac{1}{2}}J_{n+\frac{1}{2}}(\beta_{ln}a)}\left[\frac{2(2-\delta)(2n+1)(n-m)!}{3n(n+1)(n+m)!}\right]^{\frac{1}{2}},$$

$$\delta=\begin{cases}0, & m\geq 0,\\ 1, & m=0,\end{cases}\qquad V=4/3\pi a^3.$$

The components of the TM_{lmn} mode can be found from (10) by replacing the (negative) electric and magnetic field components by the analogous magnetic and electric field components, respectively, by substituting β_{ln}' for β_{ln}, and by replacing the normalization constant A by

$$A'=\frac{1}{V^{\frac{1}{2}}\left(1-\frac{(n+\frac{1}{2})^2}{(\beta_{ln}'a)^2}\right)^{\frac{1}{2}}J_{n+\frac{1}{2}}'(\beta_{ln}'a)} \times\left[\frac{2(2-\delta)(2n+1)(n-m)!}{3n(n+1)(n+m)!}\right]^{\frac{1}{2}}.$$

The general nature of spherical degeneracies is now apparent. The roots of Eqs. (7) and (8) are distinct. Hence all degenerate modes, in a particular case, must be the same type of wave, TE or TM. There is an even and odd TE or TM mode (in ϕ-variation) having the same frequency for $m>0$. Since $m\geq n$, there exist a total of $2n+1$ degenerate modes for every characteristic number. An arbitrary number of degeneracies can therefore be procurred by simply choosing the appropriate index n. For a given index l, higher values of n lead to more complex field distributions on spherical surfaces concentric with the origin.

The method which must now be employed in order to realize a desired coupling arrangement of the modes is necessarily cut-and-try. For a given number of degeneracies, the field configurations must be studied in order to determine positions at which the coupling devices may be introduced so as to establish the desired equivalent network. Usually it is simpler to employ only loops and probes at the surface of the cavity. Many possible coupling schemes exist in a spherical cavity. Several of them will be illustrated by way of example.

Figure 2 shows an arrangement in which three modes are coupled in a chain from the input to the output wave guides. The transverse magnetic field of the dominant wave-guide mode excites the TE^e_{111} spherical mode. It does not excite the TE^0_{111} mode, since the TE^e_{111} and TE^0_{111} modes have orthogonal magnetic fields at the iris. The TE_{101} mode is similarly not excited since its fields vanish at the iris. The first probe couples the TE^e_{111} and TE_{101} modes without at the same time creating spurious couplings with the TE^0_{111} mode, again by virtue of the coupling mechanism described in the preceding section. The second probe, in turn, establishes an unique coupling between the TE_{101} mode and the TE^0_{111} mode, the latter exciting a properly oriented output wave guide. Thus a three-mode chain is realized, namely,

$$\text{INPUT WAVE GUIDE}\xrightarrow{\text{iris}}TE^e_{111}\xrightarrow{\text{probe}}TE_{101}\xrightarrow{\text{probe}}TE^0_{111}\xrightarrow{\text{iris}}\text{OUTPUT WAVE GUIDE.}$$

Note that, without the presence of the coupling probes, transfer of energy to the output wave guide would not be possible because only the TE^e_{111} mode could then be set up in the cavity and the rotated output wave guide would be excited below cutoff.

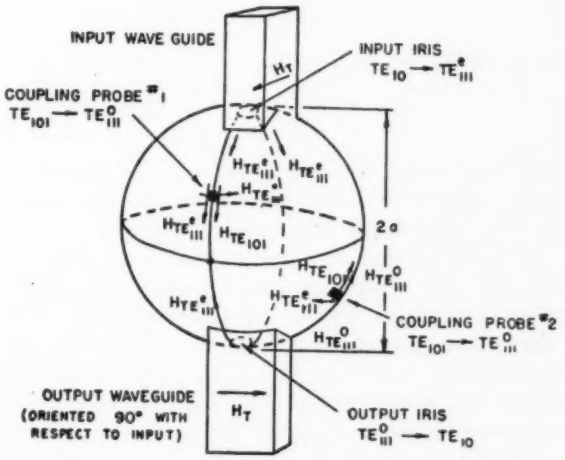


FIG. 2. Schematic drawing of the coupling mechanism of a spherical cavity excited in three modes. In the illustration, the notation under Coupling Probe #1 should read: $TE^e_{111}\rightarrow TE_{101}$.

Many other possibilities for obtaining the equivalent of three tuned circuits coupled in cascade exist. The TM_{1m1} modes can be used, in which case capacitive as well as inductive coupling is used. Higher degrees of degeneracy, in which certain of the modes are left unexcited, can also be used.

Figure 3 illustrates a scheme using both loops and probes. Five modes are coupled in cascade from input to output, giving

$$\text{INPUT}\rightarrow TM^e_{112}\rightarrow TM^0_{112}\rightarrow TM^0_{122}\rightarrow TM_{101}\rightarrow TM^e_{122}\rightarrow\text{OUTPUT.}$$

A very interesting symmetrical five-mode configuration with coupling between the first and last modes in the chain can be derived from the same type of cavity-waveguide arrangement as that shown in Fig. 2 for the three-mode cavity. By locating the coupling devices at points on the cavity wall indicated below,

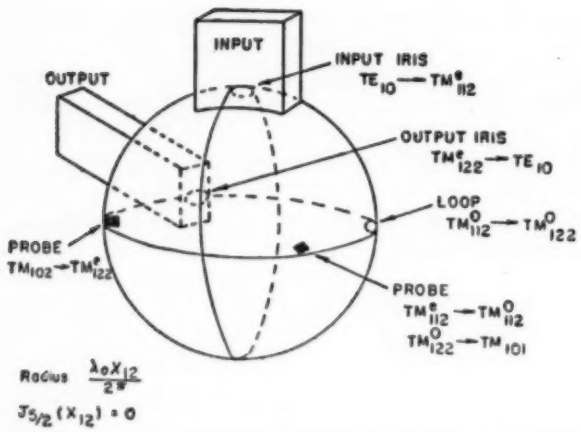


FIG. 3. Example of the coupling mechanism for a spherical cavity excited in five modes.

one obtains the scheme

$$\text{INPUT} \rightarrow TE^{0}_{112} \xrightarrow[\theta=\pi/4 \quad \phi=3\pi/2]{} TE^{e}_{122} \rightarrow TE_{102} \rightarrow TE^{0}_{122} \rightarrow TE^{e}_{112} \rightarrow$$
$$\theta=\pi \qquad \theta=\pi/2,\ \phi=\pi/4 \qquad \theta=\pi/4,\ \phi=\pi/2 \qquad \theta=\pi/4,\ \phi=\pi/4 \qquad \theta=3\pi/4,\ \phi=3\pi/2 \qquad \theta=0$$

OUTPUT.

The band-width characteristic of the structure is readily adjustable by means of the coupling between the TE_{112}^{0} and TE_{112}^{e} modes.

An investigation of higher orders of degeneracy indicates that, although an unlimited number of spherical degeneracies can be simultaneously excited, a maximum number of five can be ideally coupled in cascade. This limit coincides with that for cylindrical cavities but is here imposed by the nature of the field configurations rather than by the geometry. There are possibilities for filter configurations using more than five modes but involving cross coupling between some of the modes. Such behavior is interesting although it is unlikely to be of value in a practical device because of complexities in construction and adjustment.

In order to demonstrate the excitation of degenerate modes in a sphere, an experimental model based on the scheme shown in Fig. 2 was constructed. It was designed as an optimum-coupled filter with center frequency 9375 Mc/sec. The actual synthesis procedure employed in the design followed along the same lines as that used by Lin.[1] In the equivalent circuit, coupling coefficients were calculated from Eq. (6). Input and output coupling susceptances of the $TE^{e,0}_{111}$ modes were derived using the concept of a spherical transmission line coupled through a Bethe-hole to a uniform line. Figures 4 and 5 show the internal and external cavity construction. Two similar hemispheres were machined to close tolerances without undue effort and were then rotated 90° with respect to each other and joined under pressure on a knife edge. The probes were made adjustable by means of micrometer heads in order to facilitate the alteration of coupling between modes and to investigate the limits of the small-perturbation coupling theory.

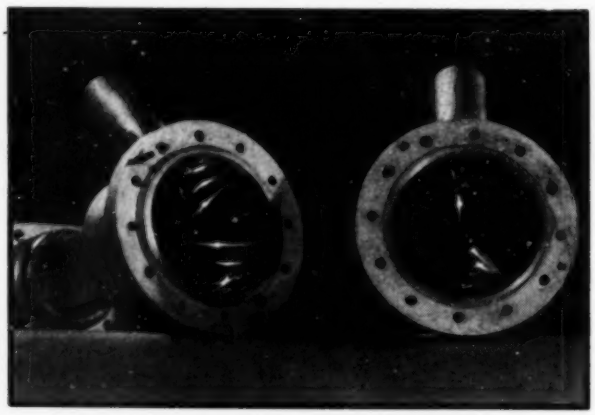


FIG. 4. Internal construction of three-mode spherical cavity. Inside diameter = 1.790 in.

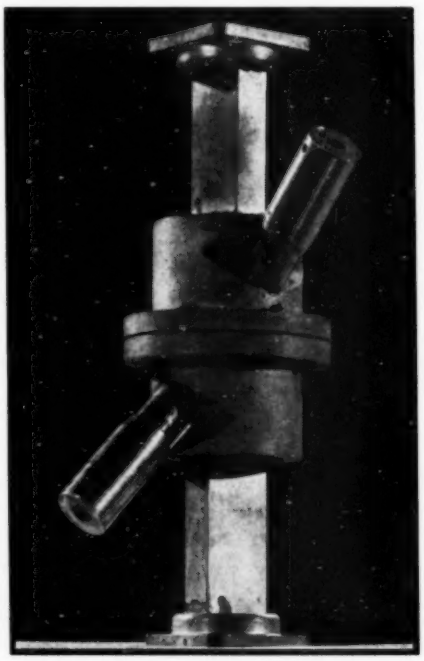


FIG. 5. External construction of three-mode spherical cavity.

Two insertion loss curves for the cavity are presented in Fig. 6. The original design was for a 10 Mc band width and a tolerance in the pass band of 1 db. The realized band width of 9 Mc and tolerance 2 db was primarily due to incidental dissipation in the cavity walls and was approximately accounted for in the theoretical curve. The couplings were redesigned for critical coupling and band width 6 Mc. The closer agreement with the theory in this case is a result of the smaller probe depths and consequently the closer compliance with the small-perturbation assumption. In fact, for optimum coupling, it is apparent that this assumption was not very closely realized in practice; the probe

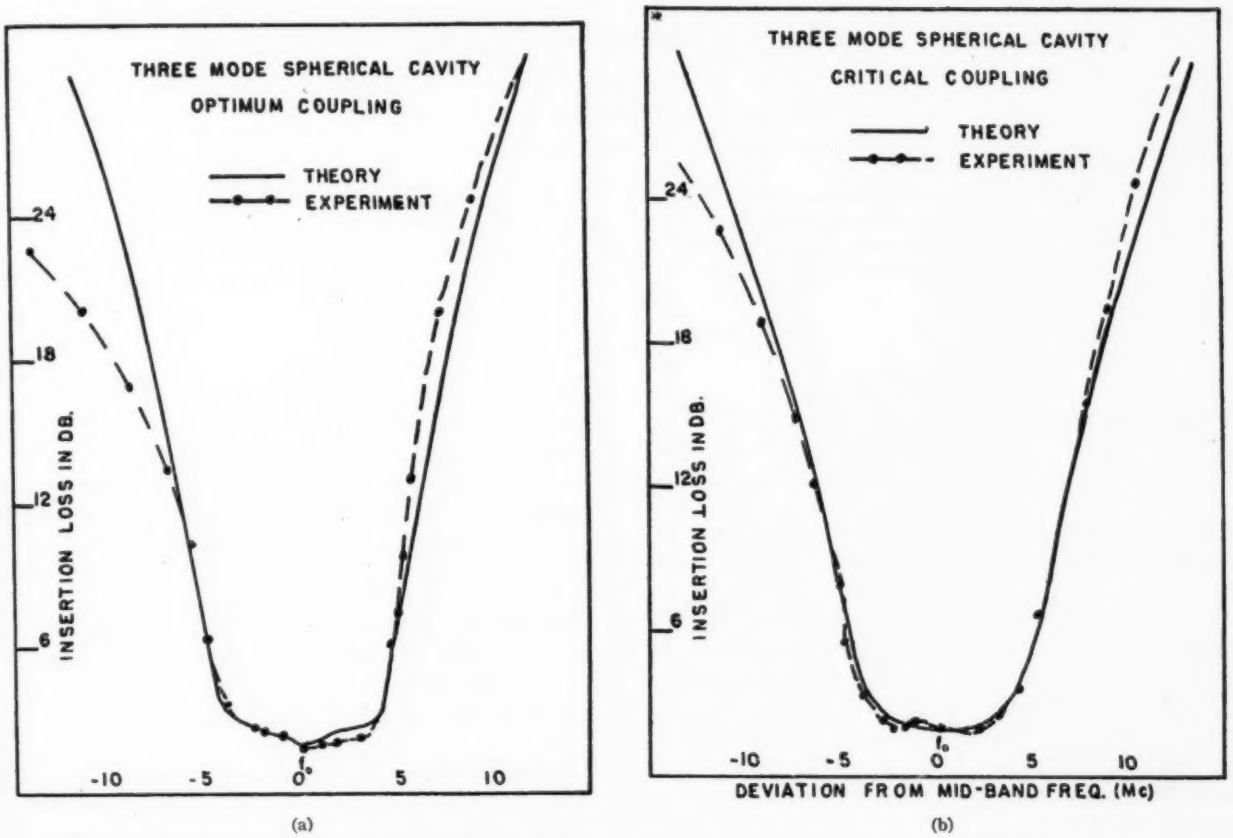


FIG. 6. Insertion-loss curves for three-mode spherical cavity. (a) Case of optimum coupling. (b) Case of critical coupling.

depth for this case was 0.235 in. For greater probe depths, a flat pass band of width up to 16 Mc was obtained but with a decreasing insertion loss on the low frequency side. In all cases the action of the cavity was exactly symmetrical with respect to input and output terminations. When the probes were adjusted exactly flush with the cavity wall, an insertion loss of greater than 55 db was measured in accordance with the theory of operation.

CONCLUSIONS

The small-perturbation technique has been extended to an arbitrary number of simultaneously excited modes having approximately the same resonant frequency. The coupling devices which were studied include irises, loops, and small volumes of conductor placed within the cavity. In each case, the physical picture of the coupling mechanism was described.

Degenerate modes in a spherical cavity have been investigated and are found to offer many interesting intermode coupling schemes. One such scheme was employed in an experimental model which served to demonstrate the technique as applied to spherical geometry. The spherical cavity was easily constructed to close mechanical tolerance and indicated the possible advantages over cylindrical cavity filters of greater mode separation and less direct coupling from input to output. Numerical designs based on coupling coefficients calculated from Eq. (6) give good agreement with theory. However, it is more convenient to regard the small-perturbation coupling theory as a guide to physically realizing a desired configuration and to then empirically adjust the intermode couplings to yield the desired behavior of the composite structure.

ACKNOWLEDGMENT

The writer wishes to express his gratitude to Dr. John R. Whinnery and Dr. W. G. Lin for their many valuable discussions and to the personnel of the Antenna Laboratory at the University of California.

The Nature of Solutions of a Rayleigh-Type Forced Vibration Equation with a Large Coefficient of Damping*

PAUL BROCK
Consolidated Engineering Corporation, Pasadena, California
(Received January 26, 1953)

In this paper, the author considers an approximation to the Rayleigh equation, $\ddot{y}+\epsilon(\frac{1}{3}\dot{y}^3-\dot{y})+y=A\cos\omega t$, and its equivalent form under the transformation $\dot{y}=v$, $y=\epsilon\xi$. A cubic function $\frac{1}{3}v^3-v+\xi$ that appears in the altered equation is replaced by a piecewise linear function. For large ϵ, the existence of periodic solutions of the approximating equation is demonstrated. Strong orbital stability properties of neighboring solutions is also shown.

1. INTRODUCTION

THE forced vibration Rayleigh equation is

$$\ddot{y}+\epsilon(\tfrac{1}{3}\dot{y}^3-\dot{y})+y=A\cos\omega t. \tag{1.1}$$

By differentiating Eq. (1.1) with respect to t and making the substitutions $\dot{y}=z$, $\omega A=-A'$, one obtains

$$\ddot{z}+\epsilon(z^2-1)\dot{z}+z=A'\sin\omega t. \tag{1.2}$$

In this form, Eq. (1.2) is a Van der Pol equation. In this paper, these equations will be restricted by the condition that ϵ be large.

The two equations have been studied extensively for the free vibration case, that is, $A=0$. Haag[1] and Dorodnicyn[2] use similar procedures in treating the free Rayleigh equation.

$$\ddot{y}+\epsilon(\tfrac{1}{3}\dot{y}^3-\dot{y})+y=0. \tag{1.3}$$

The substitutions $\dot{y}=v$, $y=\epsilon\xi$ change Eq. (1.3) to the phase plane form

$$\frac{dv}{d\xi}=-\frac{\epsilon^2}{v}\left[\frac{v^3}{3}-v+\xi\right]. \tag{1.4}$$

The curve

$$\frac{v^3}{3}-v+\xi=0, \tag{1.5}$$

is called the fundamental cubic. This curve has local extrema at the points $(\frac{2}{3}, 1)$, $(-\frac{2}{3}, -1)$ and an inflection at the origin. The integral curves of Eq. (1.4) have a zero slope on the fundamental curve. Away from the fundamental curve, the slopes of the integral curves have large absolute value.

Orders of magnitude will be represented as powers of ϵ. Thus, at a constant magnitude distance from the fundamental curve, the slope of the integral curve is of order ϵ^2. For every ϵ, the existence of a unique limit cycle has been proven.[3]

Dorodnicyn develops an expression for the limit cycle of the Van der Pol equation by considering asymptotic expansions for the solution in four basic regions of the limit curve. The expansions are developed so that they are continuations of one another in adjacent regions. The regions are depicted in Fig. 1.

Flanders and Stoker[4] prove that the limiting limit cycle for infinite ϵ consists of the vertical line segment from $(\frac{2}{3}, 1)$ to the fundamental curve, the fundamental curve itself to $(-\frac{2}{3}, -1)$, and the two symmetrical sections.

The period of the limit cycle has been proven to be $1.614\epsilon+0(\epsilon^{-1})$.[5]

Stoker[6] approximates the fundamental cubic equation (1.5) by a piecewise linear function. The resulting Rayleigh type equation then permits explicit integration and a simpler analysis is possible. In this paper, the equivalent of the fundamental cubic for the forced vibration equation has been approximated by a piecewise linear function. The solutions of the forced vibration Rayleigh type equation are then analyzed.

2. THE FORCED VIBRATION EQUATION

The substitutions $\dot{y}=v$, $y=\epsilon\xi$ change Eq. (1.1) to the form

$$\frac{dv}{d\xi}=-\frac{\epsilon^2}{v}\left[\frac{v^3}{3}-v+\left(\xi-\frac{A}{\epsilon}\cos\omega t\right)\right]. \tag{2.1}$$

Here the forcing term is seen to have a translational effect in the ξ-direction. For a significant effect, A must be of the same order of magnitude as ϵ. We introduce $\alpha=A/\epsilon$.

The frequency of the forcing term ω is chosen as $0(1/\epsilon)$; $u=\omega\epsilon$. This makes the period of the forcing term commensurate with the period of the free vibration

* Abridged from the dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at New York University, May, 1950. Presented at the International Congress of Mathematicians, September, 1950.

[1] J. Haag, Ann. Sci. Ecole Normale Superieure **60**, 35–111 (1943); **61**, 73–117 (1944).

[2] A. A. Dorodnicyn, "Asymptotic Solutions of Van der Pol's Equation" (translation from the Russian by E. Fately and P. Brock to be published by the translations project of the American Mathematical Society).

[3] N. Levinson and O. K. Smith, Duke Math. J. **9**, 382–403 (1942).

[4] D. A. Flanders and J. J. Stoker, "The Limit Case of Relaxation Oscillations," Studies in Nonlinear Vibration Theory, New York University (1946), pp. 50–64.

[5] J. LaSalle, Quart. Appl. Math. **7**, 1–19 (1949).

[6] J. J. Stoker, *Nonlinear Vibrations in Mechanical and Electrical Systems* (Interscience Publishers, Inc., New York, 1950).

response. The problem with frequency of constant magnitude has been considered by Cartwright.[7]

For forced vibrations, the analog of the fundamental curve of the free vibration equation is

$$\frac{v^3}{3}-v+\left(\xi-\alpha\cos\frac{ut}{\epsilon}\right)=0. \tag{2.2}$$

When considered in the (v, ξ) plane at any fixed time t_a, the curve is the fundamental cubic of the free vibration case displaced along the horizontal axis a distance $\alpha\cos ut_a/\epsilon$.

If t is considered to be a linearly increasing parameter, the fundamental cubic oscillates about the ordinate with frequency u/ϵ and total period $2\pi\epsilon/u$. The extreme positions assumed on the abscissa are $\pm\alpha$.

If t is considered to be a parameter, Eq. (2.2) may be approximated by piecewise linear segments. When the approximation is applied to Eq. (2.1), the differential equation is replaced by three equations of the form

$$v\frac{dv}{d\xi}+\epsilon^2\left(\frac{a}{b}v+\xi+\frac{c}{b}\right)=\epsilon^2\alpha\cos\frac{ut}{\epsilon}, \tag{2.3}$$

where a, b, and c are determined by the range of v (see Fig. 2).

Equation (2.3) may be integrated explicitly, when it is transformed into the (ξ, t) plane:

$$(\partial^2\xi/\partial t^2)+\frac{\epsilon a}{b}\dot{\xi}+\xi=\alpha\cos\frac{ut}{\epsilon}-\frac{c}{b}. \tag{2.4}$$

Although a solution is obtained for Eq. (2.4), interpretations will be made in the (v, ξ) plane (Fig. 2).

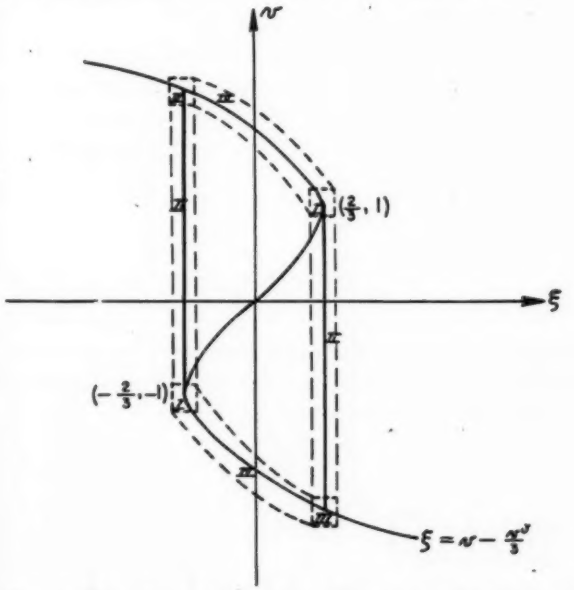


FIG. 1. The phase plane, showing the free vibration fundamental curve, the unique limiting limit cycle, and the basic regions of analysis.

[7] M. L. Cartwright, "Forced oscillations in non linear systems," Contributions to the Theory of Nonlinear Oscillations, *Annals of Mathematics Studies* (Princeton University, Princeton, 1950), XX, 149–241.

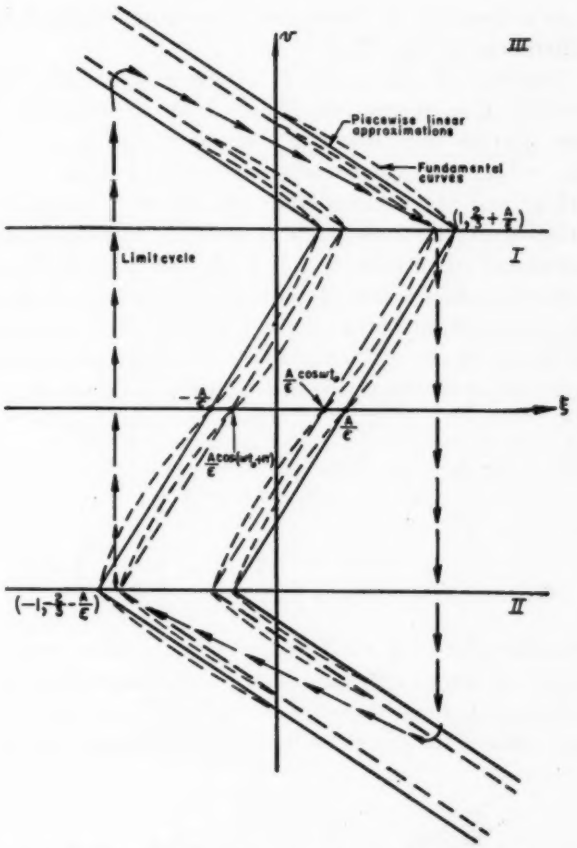


FIG. 2. The (v, ξ) plane; at time t_0, the fundamental cubic is displaced a distance $(A/\epsilon)\cos\omega t_0$. The piecewise linear approximation divides the plane into three regions (I, II, III). Also shown are the fundamental approximation curves associated with $t=0$, π, $t_0+(\pi/\omega)$. The arrow curve is a closed limit cycle corresponding to initial conditions $t=t_0$, $v=1$, $\xi=\xi_0$.

The general solution of Eq. (2.4) is

$$\xi=-\frac{c}{b}+\frac{s\left[\xi_0+\frac{c}{b}-F(t_0)\right]-[\dot{\xi}_0-\dot{F}(t_0)]}{s-r}e^{r(t-t_0)} - \frac{r\left[\xi_0+\frac{c}{b}-F(t_0)\right]-[\dot{\xi}_0-\dot{F}(t_0)]}{s-r}e^{s(t-t_0)}+F(t), \tag{2.5}$$

subject to the initial conditions, $\xi=\xi_0$, $\dot{\xi}=\dot{\xi}_0$ at $t=t_0$. r and s are the roots of the indicial equation.

$$x^2+\frac{a\epsilon}{b}x+1=0,$$

and

$$F(t)=\frac{\alpha\left[\left(1-\frac{u^2}{\epsilon^2}\right)\cos\frac{ut}{\epsilon}+\frac{au}{b}\sin\frac{ut}{\epsilon}\right]}{\left(1-\frac{u^2}{\epsilon^2}\right)^2+\left(\frac{au}{b}\right)^2}. \tag{2.6}$$

ξ as a function of t may be immediately derived by differentiating Eq. (2.5).

Starting at the point ξ_0 at time t_0 on the line $v=1(\xi=1/\epsilon)$, it may be shown from an analysis of the solution (2.5) that the time of descent to $v=-1$ $(\xi=-1/\epsilon)$ over the relaxation region I (Fig. 2) is $0(1/\epsilon)$ and the corresponding change in ξ is $0(1/\epsilon^2)$. Starting at this new point t_1 with its accompanying calculated conditions $v=-1$, $\xi=\xi_1$, the integral curve is traced across region II. This is equivalent to finding the first positive solution for $(t-t_1)$ from the ξ equation derived from Eq. (2.5), when $\xi=-1/\epsilon$. Equating lowest order terms of expansions in $1/\epsilon$, the equation

$$-1=-\bar{\eta}\exp(-k\tau)+\alpha uB^2\left[\left(\frac{u}{k}C-S\right)\cos u\tau-\left(\frac{u}{k}S+C\right)\sin u\tau\right], \quad (2.7)$$

must be solved for $\tau=(t-t_1)/\epsilon$.[8] C and S are, respectively, $\cos(ut_0/\epsilon)$ and $\sin(ut_0/\epsilon)$. k is the negative slope of the parallel line segments of regions II and III that form part of the piecewise linear approximation to the fundamental cubic.

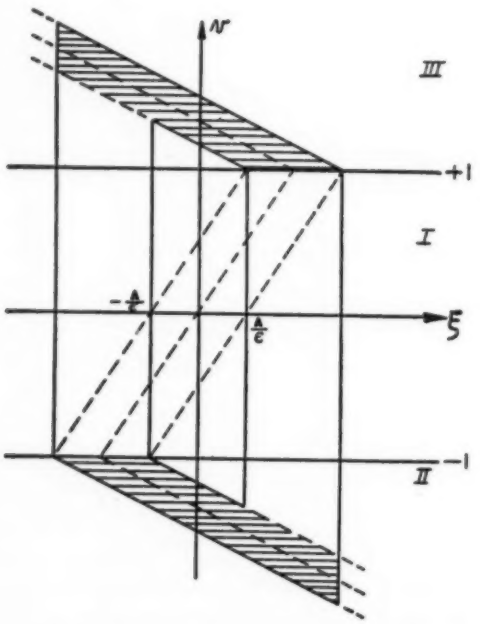


FIG. 3. Region of orbital stability of the solutions of the approximating Rayleigh equation. The two extreme fundamental curves, for $t=0$, $t=\pi$ are shown. They help to define the two parallelograms that bound the stability region. The shaded area is the region over which the integral curves are subject to a double vector field.

[8] $\tau=0$ must obviously be a solution of this equation. It can be shown that this is true if the higher order terms of expansions of $1/\epsilon$ are considered. For the first positive solution of Eq. (2.7), it is shown that the higher order terms only contribute to a higher order of magnitude error term.

$$B^2\left(1+\frac{u^2}{k^2}\right)=1,$$

$$\bar{\eta}=k\left[\xi_0+\frac{1}{k}+\frac{2}{3}-\alpha B^2\left(C+\frac{uS}{k}\right)\right].$$

Assuming $\tau>0$ is a solution of Eq. (2.7), a new point $v=-1$, $\xi=\xi_2$ may be found by evaluating Eq. (2.5). Again equating lowest order terms,

$$k\xi_2=-k\left(\frac{1}{k}+\frac{2}{3}\right)+\bar{\eta}\exp(-k\tau)\times\alpha kB^2\left[\left(\frac{u}{k}S+C\right)\cos u\tau+\left(\frac{u}{k}C-S\right)\sin u\tau\right]. \quad (2.8)$$

Adding Eqs. (2.7) and (2.8), the following result is obtained:

$$\xi_2=-\tfrac{2}{3}+\alpha\cos\left(\frac{ut_0}{\epsilon}+u\tau\right).$$

3. PERIODIC SOLUTIONS AND NEIGHBORING CURVES

The differential equation (2.3) is not altered under a half-cycle phase shift and symmetrical position shift with respect to the origin. That is, the equation is invariant under the transformation

$$\bar{v}=-v, \quad \bar{\xi}=-\xi, \quad \bar{t}=t+\frac{\pi\epsilon}{u}.$$

This permits conditions to be imposed for the existence of a periodic solution with $T=2\pi\epsilon/u$ without further tracing of the integral curve. The conditions are

$$\xi_2=-\xi_0, \quad (3.1a)$$

$$t_2-t_0=\epsilon\tau=\frac{\pi\epsilon}{u}.[9] \quad (3.1b)$$

A periodic solution will exist, therefore, if $\tau=\pi/u$ and condition (a) is satisfied. $\tau=\pi/u$ will be a solution of Eq. (2.7) under the condition

$$k\exp\left(-\frac{k\pi}{u}\right)=\frac{1-\alpha uB^2\left(\frac{u}{k}C-S\right)}{\xi_0+\frac{1}{k}+\frac{2}{3}-\alpha B^2\left(\frac{u}{k}S+C\right)}. \quad (3.2)$$

It must be shown, however, that $\tau=\pi/u$ is a first

[9] Note that the existence of subharmonic solutions would be established by imposing the alternative conditions (3.1b) $t_2-t_0=\epsilon\tau=n\pi\epsilon/u$.

positive solution. This is carried out in the author's dissertation.

Now write the conditions for the existence of a periodic solution as

$$\text{Eq. (3.2)}, \tag{3.3a}$$

$$\xi_0=\tfrac{2}{3}+\alpha C. \tag{3.3b}$$

These equations involve the amplitude and frequency of the forcing term and the initial conditions at $v=1$, $\xi=\xi_0$, $t=t_0$. If a forcing function is given, initial conditions for a periodic solution can be found. Alternatively, for a set of initial conditions, a proper forcing term may be chosen for a periodic solution.

From Eq. (2.9) the following inequality is obtained

$$-\tfrac{2}{3}-\alpha\leq\xi_2\leq-\tfrac{2}{3}+\alpha. \tag{3.4}$$

This shows that the value of ξ_2 on the line $v=-1$ is bounded independently of the initial conditions. This property may be generalized, and it can be shown that all solutions lie in the ring-like region bounded by the parallelograms shown in Fig. 3, with the possible exception of their initial segments. This is a direct analog of the strong stability demonstrated in the free vibration case.

More can be said about neighboring solutions to a periodic solution. A neighboring solution will parallel the periodic solution in region I, cross it in region II, and parallel it again in region I (see Fig. 4). The analysis of this paper is not sufficiently detailed to follow the integral curve further. An interesting point, however, is that integral curves may intersect each other and also may be self-intersecting. This divergence from free vibration theory is due to a double vector field determined by the differential equation in regions II and III.[10] It can be seen that in the shaded regions of Fig. 3, any integral curve must follow the progressing fundamental curve. In particular, within the shaded regions of Fig. 3, at any time t, the integral curve must be within $0(1/\epsilon^2)$ of the curve

$$-\frac{a}{b}v+\xi+\frac{c}{b}-\alpha\cos\frac{ut}{\epsilon}=0. \tag{3.5}$$

At all times,

$$v=\epsilon\frac{d\xi}{dt}; \tag{3.6}$$

[10] The Poincaré-Bendixson theorem, applicable to the free vibration case is based on the fact that the integral curve lies in a region over which the differential equation defines a unique vector field.

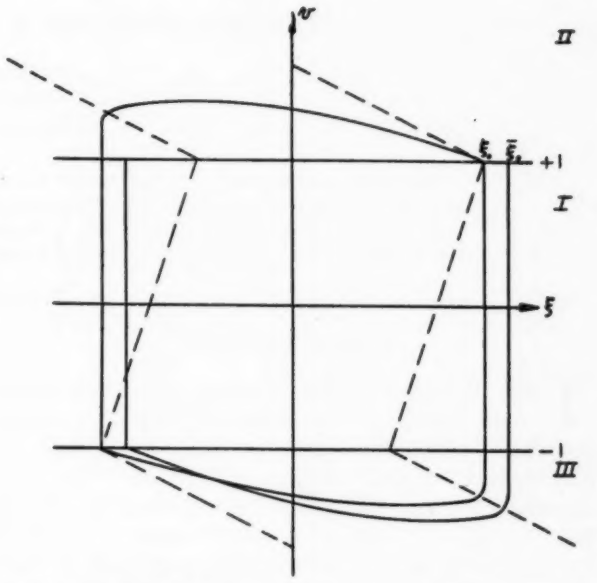


FIG. 4. Periodic solution and neighboring integral curve. The periodic solution has an initial condition at $t=t_0$ of $\xi=\xi_0$, $v=+1$; the neighboring solution starts at $\xi=\bar{\xi}_0$, $v=+1$ at the same time, $t=t_0$.

hence, the tangent of the integral curve at any point as a function of time is

$$\frac{dv}{d\xi}=-\frac{b}{a}\left(\frac{\alpha u}{v}\sin\frac{ut}{\epsilon}+1\right). \tag{3.7}$$

(Note that these equations are lowest order term approximations in $1/\epsilon$.) At any point of the shaded regions, therefore, the integral curve considered as a vector function is double valued, depending on whether the fundamental curve associated with a point is considered at time t_α or $2\pi-t_\alpha$.

Qualitative verification of these results was obtained by solving the problem on the Reeves Electronic Analogue Computer (REAC). The machine allows one to easily vary the parameters ϵ, A, ω, and the initial conditions. Hence, it was possible to solve a large number of problems to verify the results.

The author wants to express his sincere appreciation to Professor Fritz John and Professor J. J. Stoker of New York University for their guidance, assistance, and encouragement. He also thanks Dr. Stanley Fifer of Reeves Instrument Corporation for his valuable cooperation.

Diffusion of Barium in an Oxide-Coated Cathode*

RENATE S. BEVER
Cornell University, Ithaca, New York
(Received February 27, 1953)

The diffusion of Ba through activated BaO-SrO cathode coatings was studied as a function of temperature. Most of the radioactive tracer material diffused according to a dependence of $\log_{10}D$ *vs* $1/T$ similar to that obtained by Redington for diffusion of Ba in single crystals of BaO. The activation energy was 4.1 ± 0.6 ev above 1280°K and 0.40 ± 0.07 ev below that temperature.

I. INTRODUCTION

THE diffusion of Ba through the oxide-coated cathodes plays an important part in several theories[1–6] which have been proposed to explain the performance of such cathodes. Activation has been thought to occur through surface adsorption of free Ba, through an increase of stoichiometric excess of Ba in the body of the coating, or through both of these processes. Diffusion of free Ba from the core metal, where it is liberated, is required as the Ba distributing mechanism in the above theories.

Blewett[3] has calculated a temperature dependence of the diffusion process of 0.7 ev between 820°K and 1050°K from a study of emission decay characteristics as a function of temperature. Claasen and Veenemans[7] observed that between 1300°K and 1500°K the vapor pressure and speed of diffusion of BaO in BaO-SrO mixtures were proportional to one another. The present investigation was carried out to supply additional quantitative information on the Ba diffusion in BaO-SrO coated cathodes, comparable to Redington's[8] study of Ba diffusion in single crystals of BaO.

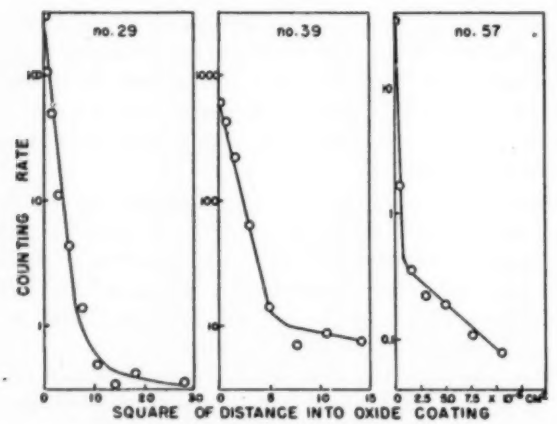


FIG. 1. Counting rate in scales of 64 per minute *versus* the square of the distance into the oxide coating. No. 29 was held at $1000/T = 1.30$ for 4100 minutes, No. 39 at $1000/T = 0.705$ for 28 minutes. No. 57 was a control run; it was sectioned after evaporation without being given any diffusion time. Background for No. 29 was 1.20; for No. 39, 1.60; and for No. 57, 0.32 scales of 64 per minute.

II. EXPERIMENTAL PROCEDURE

A mixture of 47 percent $BaCO_3$, 53 percent $SrCO_3$ in nitrocellulose binder was sprayed on the flat circular ends of $\frac{1}{8}$-in. diameter nickel "end-fire" cathodes. The average carbonate particle size was 1.5μ; the coating thickness was about 0.01 cm and its density was 1.3 ± 0.1 g/cm³. The surface of the coating was sliced until it was parallel to the edge of the knife of the microtome, which has been described by Redington.[8] The portion of the coating near the intersection of the cylindrical and flat circular surfaces of the nickel sleeve was removed.

The cathode was mounted into the evaporator assembly adapted from Redington's[8] design: The end-fire cathode and its heater replaced the crystal mount, and a thermocouple junction was spotwelded on the nickel sleeve as close to the carbonate coating as possible. Under a pressure not exceeding 10^{-4} mm Hg, the evaporator cup containing the compound with tracer Ba^{140} was outgassed by slowly heating it to about 1350°K. The mixture of $BaCO_3$ and $SrCO_3$ of the cathode was converted to the oxides by slowly heating the cathode to about 1150°K. The cathode was activated by heating it at 1400°K for a few minutes while drawing emission current. It was then allowed to cool, and the BaO containing the Ba^{140} was evaporated onto its emitting surface by heating the evaporator cup to between 1400°K and 1500°K for approximately one half-hour. The cathode was then reheated for 30 seconds at 1400°K while drawing emission current.[9] It was brought to the temperature at which diffusion was to be studied and maintained there for a time ranging in order of magnitude from 10 minutes at 1500°K to 100 hours at 750°K. During this interval the evaporator cup was usually kept at approximately the same temperature as the cathode to reduce net evaporation of tracer BaO and minimize heat radiation from the cathode surface.

* This work has received support from the U. S. Office of Naval Research.

[1] J. A. Becker and R. W. Sears, Phys. Rev. **38**, 2193 (1931).
[2] J. P. Blewett, J. Appl. Phys. **10**, 688 and 831 (1939).
[3] J. P. Blewett, Phys. Rev. **55**, 713 (1939).
[4] R. L. Sproull, Phys. Rev. **67**, 166 (1945).
[5] A. S. Eisenstein, *Oxide Coated Cathodes, Advances in Electronics I* (Academic Press Inc., New York, 1948).
[6] L. S. Nergaard, RCA Rev. **13**, 464 (1952).
[7] A. Claasen and C. F. Veenemans, Z. Physik **80**, 342 (1933).
[8] R. W. Redington, Phys. Rev. **87**, 1066 (1952).
[9] This caused some premature diffusion, but the reheating was necessary to compensate for loss in activation during the evaporation of the tracer BaO. The emission curent of the cathodes thus treated was usually about 100 ma/cm² at 1100°K.

Although no evidence was gathered from this work of the dependence of the diffusion constant D on the state of activation, the above procedure was undertaken to simulate an operating cathode.

The cathodes had to be exposed to room air for sectioning. About one minute after this exposure many of the oxide coatings peeled from the nickel base. This peeling seemed to be independent of the treatment of the nickel base, but was influenced by the humidity conditions at the time of spraying and sectioning. Because of this difficulty the coating thickness could not be much more than 0.01 cm. No measurements were made on "peeled" coatings.

The cathode coatings were trimmed at the edges, sectioned in the microtome, and the radioactive counting rates obtained according to the procedure outlined by Redington.[8] It was usually possible to cut from 12 to 16 sections, each 5×10^{-4} cm thick. Alignment in the microtome was generally such that, beginning with the third slice, complete cross sections of the coating were cut.

III. EXPERIMENTAL RESULTS

Typical plots of counting rates C *versus* the square of the distance x^2, at which the sections were cut are reproduced in Fig. 1. No. 29 is representative of distribution curves below 1280°K and No. 39 above 1280°K. No. 57 shows the curves for a control cathode coating which was treated as all others except for not being given any diffusion time t. During the hour of evaporation of tracer BaO, its temperature was 420°K. The plots demonstrate that most of the tracer BaO diffused by a relatively slow process. However, a few percent of tracer BaO is found deeper in the coating, as evidenced by the "tails" with low counting rates.

To obtain the diffusion constant D for the process by which most of the material moved the "tail" was treated as background. D was calculated from Eq. (1) below, which is discussed by Redington.[8] It can be derived from Fick's law for a semi-infinite solid with the boundary condition of a given quantity of tracer material Q deposited at the origin,

$$C=Q(\pi Dt)^{-\frac{1}{2}}\exp(-x^2/4Dt). \quad (1)$$

The dependence of D on the reciprocal of absolute temperature $1/T$ is shown in Fig. 2. In the range of our measurements D lay between $.3\times10^{-9}$ cm²/sec at 1490°K and 1.5×10^{-12} cm²/sec at 740°K. The temperature dependence can be described by

$$D=D_1\exp(-E_1/kT)+D_2\exp(-E_2/kT). \quad (2)$$

The activation energy E_1 above 1280°K is 4.1 ± 0.6 ev; below 1280°K E_2 is 0.40 ± 0.07 ev. The errors are computed at the 5 percent confidence level.[10] D_1 is 10^5 cm²/sec and D_2 is 4.5×10^{-10} cm²/sec.

The "tail" distribution mentioned previously appeared even in the control run No. 57, although it fell off more steeply in No. 57 than in all other runs.

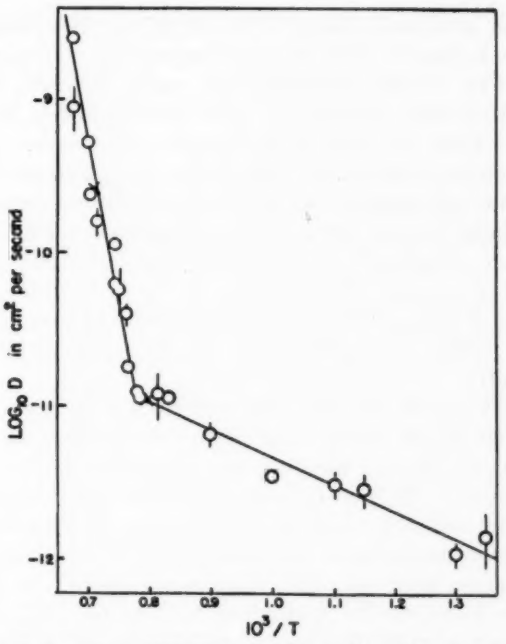


FIG. 2. $\log_{10}D$ *vs* $1000/T$ for oxide coatings. The circles show data points obtained on coatings sprayed as described in the text; the crosses are for commercially prepared RCA "end fire" cathodes type FK 120 R2-602A. The vertical lines indicate probable error in the measurements.

The "tails" might have been caused by fast grain boundary diffusion which in No. 57 occurred during the hour of tracer BaO evaporation. However, it is felt that the "tails" could equally well be an experimental artifact.

IV. DISCUSSION AND APPLICATION TO OXIDE-COATED CATHODES

The superficial appearance of the $\log_{10}D$ *versus* $10^3/T$ curve is like that found by Redington[8] for the neutral diffusion in single crystals of BaO. There are, however, two differences.

(1) Only one diffusion process was found in the high-temperature region (apart from the "tail" portion). At $10^3/T=0.75$, two runs were made with the diffusion time differing by a factor of 10, and both distribution curves showed only the steep, "near-the-surface" portion and the "tail" portion. It is possible, but not very probable, that a third slope was missed because of the limited thickness of the cathods coatings and the scatter of the points on the distribution curves.

(2) The temperature dependence in the high-temperature region is 4.1 ev, not 11.0 ev. The Dt products for most of the points in Fig. 2 did not vary by more than a factor of two, and so it is not likely that the highest-temperature points belong to a slower process. Since nickel was used as the cathode base metal contamination of the thermocouple junction was a possibility, but Vycor tubing was used to prevent this, and several calibrated thermocouples were employed.[11]

[10] H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1946), p. 548 ff.

[11] The question whether the data shown in Fig. 2 represents grain boundary diffusion as compared to Redington's (see reference

The activation energy of 4.1 ev in the high-temperature region is close to the temperature dependence of 3.5 ev of the vapor-pressure curve for BaO in a 50-50 mixture of BaO and SrO.[7] This is in fair agreement with the observation made by Claasen and Veenemans,[7] mentioned previously. To ascertain that one was not dealing with a "pseudo" diffusion constant simulated by pore evaporation of BaO, the magnitude of such a "pseudo" D was estimated but gave results too small by a factor of 300.

The 0.40 ev in the low-temperature region is of the same order of magnitude as the 0.7 ev obtained by Blewett.[4]

Diffusion of Ba plays an important role in several theories of the oxide-coated cathode. In applying the results of Fig. 2 to this problem we are faced with two limitations. (1) The D measured is for the motion of barium, but a barium excess can also be created or moved in the cathode coating by the diffusion of oxygen vacancies; such vacancies have been shown[12] to exist and to move readily in single crystals of BaO. (2) It is not known from the present experiments what charge (if any) the diffusing Ba carries. It seems possible from Redington's[8] measurements that we are dealing here with the diffusion of positive ions carrying 1.7 ± 0.3 electronic charges. If it were known that the Ba^{140} was so charged, one could use the Einstein relation to determine the electrolytic mobility. Attempts to apply an electric field to the samples in the present experiment failed because the means for applying it were inconsistent with good diffusion constant measurements.

It is interesting, however, to examine the order of magnitude of the effects produced in a cathode by barium diffusing. The exact solution of the problem of diffusion through a cathode of finite thickness, with a supply of free Ba at the core boundary and a removal or Ba at the emitting surface by evaporation and electrolysis, is fairly complex and is not done here. However, an idea of the order of magnitude of the diffusion currents I can be obtained by solving representative problems for the semi-infinite case.[13,14] For instance, if one assumes a constant concentration of one monolayer of Ba (10^{15} Ba atoms/cm^2/monolayer) maintained at the core surface and uses the measured D of 6×10^{-10} cm^2/sec at 1425°K (a typical "activation" temperature) one can estimate the diffusion current arriving at the vacuum interface of the cathode. If the cathode is 0.002 cm thick (a typical cathode thickness) one finds at time $T=1$ minute a diffusion current of 2×10^{-9} monolayer/min, at $T=5$ minutes $I=5$ monolayer/min, and at $T=9$ minutes $I=50$ monolayer/min. The evaporation rate of BaO from an oxide cathode of equal parts of BaO and SrO at 1425°K is 55 monolayers/min.[7] (This decreases to about 30 monolayers/min after 10 minutes.[7]) The example indicates that, until the steady state of flow of Ba is reached, considerable depletion of BaO from the surface layers of an oxide cathode occurs during activation. This is in agreement with the results of electron diffraction experiments.[5]

Continuing with the above example, one may now consider what happens as the temperature of the cathode is lowered to 1000°K (representative of operating temperatures). The product of the Ba concentration gradient established at 1425°K (1.4×10^{24} atoms/cm^3/cm after $T=9$ minutes) and the measured D at 1000°K (7×10^{-12} cm^2/sec) gives, immediately after lowering the temperature, a rate of Ba arrival of 0.6 monolayer/min by diffusion. Because of evaporation, the rate of departure of BaO from the emitting surface at 1000°K is only 10^{-4} monolayer/min.[7] Thus a replenishing of surface Ba could occur were it not for removal of Ba by electrolysis during the drawing of emission current. A crude calculation of the Ba ion flow leaving the emitting surface by electrolysis gives an order of magnitude of one monolayer/min at an emission current of 0.1 amp/cm^2; this is calculated from previous estimates.[15] Much more careful calculations would be necessary to find in what time the steady state in the diffusion-electrolysis process could be reached and thus to decide whether electrolytic explanations[1,4,6] of emission decay phenomena are compatible with the magnitude of the diffusion constant of Ba.

It is felt that diffusion of Ba in oxide-coated cathodes is sufficiently rapid to maintain a distribution of excess Ba in the body of the coating. This agrees with semiconductor theories of high oxide-cathode activity.[5] The crude calculations above seem to indicate that the diffusion currents are not large enough to completely replenish the evaporation loss of Ba from the emitting surface. This is not compatible with the theory[1] which explains cathode activation on the basis of adsorption of a polarized surface layer of Ba.

The author is greatly indebted to Professor R. L. Sproull for suggestion of and assistance with this work and to Dr. R. W. Redington for the use of his apparatus.

8) bulk diffusion is not settled here. Redington estimated that grain boundary diffusion would be the predominant process for oxide particles smaller than 0.2μ. However, the oxide crystalite size was not measured in the present work, and measurements by other authors are conflicting. A. S. Eisenstein [J. Appl. Phys. **17**, 434 (1946)] using the line broadening of x-rays, found that for an oxide coating heated for 5 minutes at 1325°K the crystal size was greater than 10μ. J. Shimazu, [J. Phys. Soc. Japan **6**, 479 (1951)] using the small angle scattering of x-rays, found that for an oxide coating heated for 15 minutes at 1370°K, the crystal size was about 0.035μ. These values straddle the estimated critical size of 0.2μ. The latter estimation is in itself somewhat doubtful because of the complexity of the theoretical problem of grain boundary diffusion, as discussed by J. C. Fisher [J. Appl. Phys. **22**, 74 (1951)] and others.

[12] R. L. Sproull and R. S. Bever, Phys. Rev. **90**, (1953). A more detailed account of this work is being prepared for publication.

[13] H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids* (Clarendon Press, Oxford, 1947), Chaps. II and III.

[14] In these calculations the "tails" are again considered as background, and the D measured for the slower process, plotted in Fig. 2, is used.

[15] J. A. Becker, Phys. Rev. **34**, 1323 (1929) indicates that in BaO at 800°K the fraction of the total current carried by the Ba ions is 4×10^{-3}; T. P. Berdennikowa, Physik. Z. Sowjetunion **2**, 77 (1932) estimates this fraction to be 5×10^{-5} at 1273°K. In the present calculation, at 1000°K in a 50–50 mixture of BaO and SrO, this fraction is assumed to be about 1×10^{-4}.

Degradation of Polyisobutylenes on Shearing in Solution*†

A. B. BESTUL AND H. V. BELCHER
National Bureau of Standards, Washington, D. C.
(Received March 20, 1953)‡

The decrease of viscosity on shearing in a capillary at 37.8°C and a nominal rate of shear of 65 000 sec^{-1} has been measured for 5, 10, and 15 weight-percent solutions in *o*-dichlorobenzene of 10 polyisobutylenes having viscosity-average molecular weights from 40 000 to 2 300 000. Very little, if any, viscosity decrease was observed for polymers having molecular weights below 500 000, but marked decreases were observed for polymers with molecular weights above 500 000. The observed decreases were not recovered on standing and correspond to appreciable decreases in the molecular weights of the polymers. The molecular weight reduction can be explained on the basis of polymer chain rupture if the mechanical energy developed in a volume less than that enveloping a single polymer molecule can be concentrated into a single bond.

I. INTRODUCTION

WHEN polymer solutions are sheared their viscosities often decrease as the shearing proceeds.[1–3] This effect is usually thought to be the result of the rupture of the polymer molecule chains.[4,5] The dependence of this decrease on the size of the polymer molecules seems important. This phenomenon has therefore been investigated in 5, 10, and 15 weight-percent solutions, in *o*-dichlorobenzene, of 10 different unfractionated polyisobutylenes having viscosity average molecular weights from 40 000 to 2 300 000.

II. EXPERIMENTAL MATERIALS

The polyisobutylenes investigated were the commercial Vistanexes, produced by the Enjay Company, Inc. Table I lists these polymers by their commercial designations and shows the mid-points of the ranges given by the manufacturer for the viscosity-average molecular weights M of the different grades. This information was originally given in the form of Staudinger values, but was recalculated using the equation

$$[\eta]=KM^{\alpha}, \qquad (1)$$

where K and α are constants, having values of 3.6×10^{-4} and 0.64, respectively, in di-isobutylene at 20°C. The corresponding intrinsic viscosities for these conditions are also given.

The *o*-dichlorobenzene used was received from the University of Akron, Government Laboratories without indication concerning original source.

Solutions of each polymer in *o*-dichlorobenzene were prepared at concentrations of 5, 10, and 15 weight-percent by placing weighed quantities of polymer and solvent into 8-ounce cylindrical glass jars and rolling them until the mixtures appeared homogeneous to the naked eye. No purification was performed on the polymers, solvent, or solutions.

III. APPARATUS

The present investigations were made principally with the McKee Worker-Consistometer.[6] In the measurements reported here the experimental material is sheared by being forced through a 0.2-mm diameter capillary in a 0.25-inch thick steel disk. This disk is clamped between two coaxial cylindrical tubes of steel (0.5 in. i.d.), each of which is equipped with a close fitting piston. The experimental material is placed between the disk and the piston in one of the cylindrical tubes. The piston can then be used to force the material through the shearing disk, into the opposed cylinder. Since the instrument is essentially symmetrical with respect to the cylinders and pistons on the two sides of the disk successive measurements can be made by moving the pistons in alternate directions. The assembly of disk, cylinders, and pistons is housed in a thermo-

TABLE I. Polymers investigated.

Vistanex	$[\eta]$,[a] dl/g di-isobutylene, 20°C	Visc. Av. Mol. Wt.[b]
LM-S	0.315	39 000
LM-MS	0.355	48 000
LM-MH	0.405	58 000
LM-H	>0.43	>64 000
B-40	1.20	300 000
B-60	1.80	600 000
B-80	2.40	1 000 000
B-100	3.00	1 300 000
B-120	3.60	1 800 000
B-140	4.20	2 300 000

[a] Intrinsic viscosity, $[\eta]=\lim_{c\to0}\left(\frac{1}{c}\right)\left(\frac{\text{viscosity of solution}}{\text{viscosity of solvent}}-1\right)$.
[b] In Eq. (1), $K=3.6\times10^{-4}$ and $\alpha=0.64$.

* The work reported here was financed by the Reconstruction Finance Corporation, Synthetic Rubber Division.
† Presented before the 299th Meeting of the American Physical Society, Washington, D. C., April 27–29, 1950.
‡ Original unrevised manuscript was received October 31, 1952.
[1] H. Willenberg and W. Fritz, Angew. Chem. **B20**, 134 (1948).
[2] R. Schnurmann, Proc. International Rheol. Congress, Holland, 1948, North Holland Publishing Company (Amsterdam), 1949, p. II–142.
[3] Fenske, Klaus, and Dannenbrick, Report to the American Petroleum Institute, 1949.
[4] C. N. Davies, Proc. International Rheol. Congress, Holland, 1948, North Holland Publishing Co. (Amsterdam), 1949, p. II–152.
[5] H. W. Melville, Science Progr. **38**, 1 (1950).
[6] S. A. McKee and H. S. White, ASTM Bull. No. 153, 90 (1948); also J. Research Natl. Bur. Standards **46**, 18 (1951).

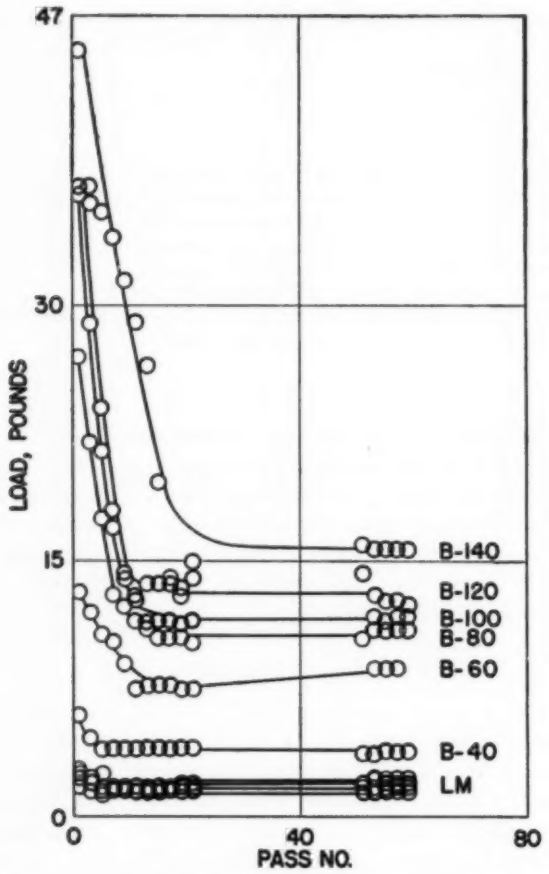


FIG. 1. Shearing load reductions with continued shearing for five percent solutions of various molecular weight polyisobutylenes in *o*-dichlorobenzene.

stated air bath. The pistons are driven back and forth by an eccentric cam operating through a yoke, and the force necessary to hold the disk and cylinders stationary is registered on a scale.

IV. MEASUREMENTS, RESULTS, AND DISCUSSION

A. Viscosity Decrease

Each solution investigated was put into the consistometer; the instrument thermostated at 37.8±0.5°C, and operated at a nominal rate of shear of 65 000 sec^{-1}, if nominal rate of shear is used as the quantity obtained from the equation which gives the rate of shear at the capillary wall in Newtonian flow. This rate corresponds to an average linear flow velocity in the capillary of 2.7 m/sec. The load readings were then recorded for every other pass of the solution through the capillary for about 50 passes. Since the nominal rate of shear was maintained constant the observed load was proportional to the apparent viscosity, the proportionality factor being apparent viscosity in poises equals 0.08 times load in pounds.

The results of all the measurements at the extreme concentrations (5 and 15 percent) are shown in Figs. 1 and 2, where the abscissas represent the successive passes of a solution through the capillary and the ordinates represent the load readings for these passes. The 15 percent solutions of Vistanex B-100, B-120, and B-140 were too viscous to handle in the present apparatus, therefore no data were obtained for these solutions. The most noticeable feature of both these figures is that for the lower molecular weight polymers only small, if any, viscosity reductions are shown, while for the higher molecular weight polymers pronounced viscosity reductions appear. The results for 10 percent solutions form plots very similar to those shown for the 5 and 15 percent solutions, the magnitudes of the values lying between those for the extreme concentrations. In Fig. 3 the viscosity reductions, in percents of original viscosities, are plotted against the molecular weights of the polymers. The values for the solutions of the *LM* series polymers were obtained from plots having more sensitive scales than Figs. 1 and 2. In view of the low load readings for the *LM* solutions (see Figs. 1 and 2), the viscosity reductions shown in Fig. 3 for these solutions are so uncertain that their very existences are not demonstrated with certainty. Uncertainties of the same source also apply, to a lesser degree, to the B-40 solu-

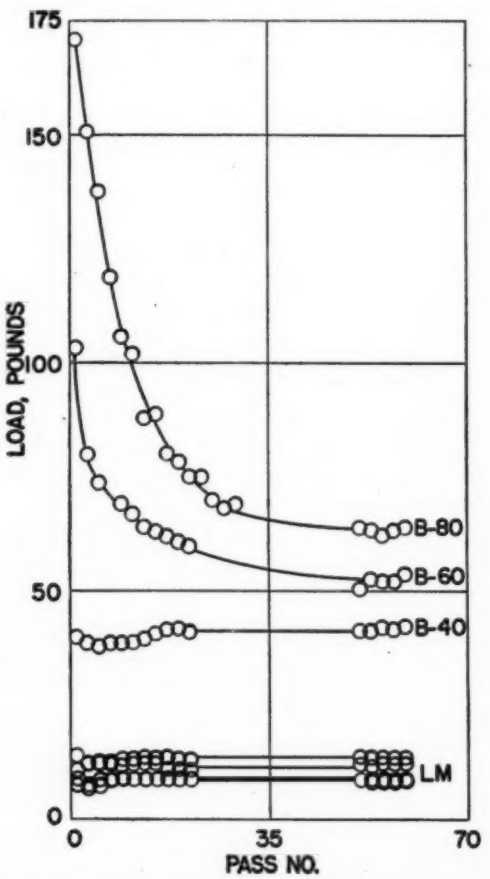


FIG. 2. Shearing load reductions with continued shearing for fifteen percent solutions of various molecular weight polyisobutylenes in *o*-dichlorobenzene.

tions of lower concentrations. Consequently the data in Fig. 3 are not precise enough to yield a quantitative relation between viscosity reduction and the molecular weight of the polymer. They do, however, show clearly that proportionately larger viscosity reductions occur for the higher molecular weights than for the lower. The data do not indicate whether the differences in viscosity reduction with concentration of solution observed for the individual polymers are significant. It can be concluded from the data shown, however, that any concentration dependence must be much less than the molecular weight dependence.

B. Apparent Recovery

If a solution which shows a viscosity decrease is allowed to rest for a time after being passed 50 times through the capillary, and then measurements resumed, the viscosity observed immediately after the resumption is higher than that last observed. As the resumed shearing continues, however, the viscosity again decreases to the value observed just before the rest. There was some question as to whether this apparent recovery might be an inherent property of the solutions, or an instrumental effect. If it were an inherent property of the solution, the part of the viscosity decrease which is recovered would be associated with thixotropy. To investigate this question, 10 percent solutions of B-120 were subjected to measurements identical with the above described ones except that no shearing element was included in the instrument assembly. Under these conditions, the force observed is that which originates only in the clearance space between the cylinders and pistons. With this arrangement there occurred apparent recoveries of magnitude equal to those mentioned above. This demonstrates that the apparent recoveries are not associated with the flow properties of the solutions but probably arise from evaporation of solvent from the film exposed on a cylinder's inner wall during the rest period.

TABLE II. Effect of shearing on dilute solution viscosities Vistanex B-120 in *o*-dichlorobenzene, 37.7°C.

History	Intrinsic viscosity $[\eta]$, dl/g	k' constant
Unsheared	1.95	1.0
Sheared	1.55	1.4

C. Dilute Solution Viscosity

To check the supposition that the observed viscosity decreases resulted from the rupture of molecular chains, viscosities were measured for dilute (0.25 and 0.5 weight-percent) solutions prepared from the 10 percent solution of B-120, both before and after shearing for 50 passes. These measurements were made in a Bingham viscometer (0.0109 cm capillary radius, 12 cm capillary length) at 37.7°C. Dilute solutions were prepared from the sheared sample both immediately after shearing and after letting the sample rest a day after shearing. The results for the two procedures were the same, again showing that no real recovery occurs. The results for the sheared and the unsheared samples in terms of intrinsic viscosities and k' constants of the equation

$$\eta_{sp}/c=[\eta]+k'[\eta]^2c \qquad (2)$$

are given in Table II. Since the intrinsic viscosity of this polymer in di-isobutylene, a good solvent, is 3.6 dl/g (see Table I), and since k' for a linear polymer in a good solvent is usually about 0.3 to 0.4,[7] one can conclude that *o*-dichlorobenzene is not a good solvent for polyisobutylene. No significance is attributed to the difference appearing in k' for the sheared and unsheared samples.

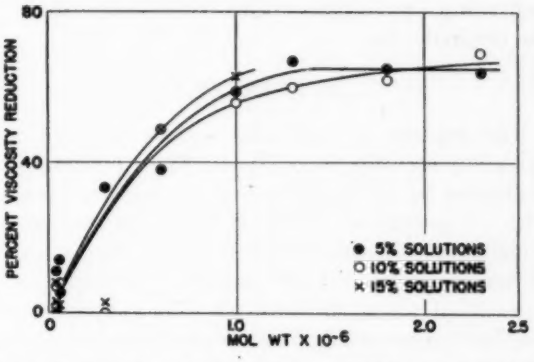


FIG. 3. Percent viscosity reductions at 65 000 sec^{-1} for solutions of various molecular weight polyisobutylenes in *o*-dichlorobenzene at three concentrations.

The smaller intrinsic viscosity for the sheared material (to about a factor of 0.8) indicates that the observed viscosity decrease is associated with an appreciable decrease in molecular weight and substantiates the occurrence of molecular chain rupture. However, it is most informative to evaluate the amount of molecular chain rupture quantitatively in terms of molecular weight decrease. Since the values for the constants K and α in Eq. (1) are not known for the polyisobutylene-*o*-dichlorobenzene system, no absolute values for molecular weights can be calculated from the above intrinsic viscosities. However, since α is often about $\frac{2}{3}$ for solutions of polymers having linear flexible molecules,[8] a probable value for the ratio of the molecular weights of the sheared and unsheared polymers may be obtained by using $\alpha=\frac{2}{3}$. One thereby obtains

$$\frac{M_{\text{sheared}}}{M_{\text{unsheared}}}=\left(\frac{[\eta]_s}{[\eta]_u}\right)^{\frac{1}{2}}=\left(\frac{1.55}{1.95}\right)^{\frac{1}{2}}=0.7,$$

[7] C. E. H. Bawn, *The Chemistry of High Polymers* (Interscience Publishers, Inc., New York, 1948), p. 167.

[8] H. Mark and A. V. Tobolsky, *Physical Chemistry of High Polymeric Systems* (Interscience Publishers, Inc., New York, 1950), p. 289.

indicating a molecular weight reduction to about 0.7 of the original value.

V. ENERGETICS

The rupture of molecular chains must be accompanied by certain energy effects. These effects should be evaluated to determine whether they are compatible with the energy available. If the chain rupture in the present case consists of the breaking of aliphatic carbon-carbon bonds and the compensatory formation of other unspecified bonds, the most plausible conjectures about possible compensatory bonds likely to be formed suggest that the over-all energy requirement is such that no energy need be added to the system. However, the chain rupture can be considered as a chemical reaction in which at least part of the activation energy is supplied mechanically rather than thermally. Therefore one should investigate whether the mechanical energy available is sufficient to supply the required activation energy. This question cannot be investigated exactly since an undetermined part of the activation energy is supplied thermally rather than mechanically, and an undetermined part of the mechanical energy supplied is probably dissipated immediately as heat to overcome internal friction and cannot be concentrated in space as is found necessary in what follows. However, it is informative to compare the total activation energy with the total mechanical energy supplied.

The most reasonable value to assume for the activation energy in this case is probably 80 kcal or 33.5×10^{11} ergs per mole of broken bonds, the bond energy of aliphatic carbon-carbon bonds.[9] This is equivalent to 5.5×10^{-12} ergs per broken bond. The energy dissipated per unit volume when material is forced through a capillary is given by the pressure drop across the capillary. A load of 50 pounds corresponds here to a pressure drop of 250 psi or 17.2×10^{6} dynes/cm^2 $=17.2\times10^{6}$ ergs/cm^3. Therefore the volume in the flow field within which 5.5×10^{-12} ergs are developed in passing through the capillary is 3.2×10^{-19} cm^3. The volume occupied by two methylene groups connected by a covalent bond is around 10^{-23} cm^3. There could not be developed in this volume sufficient activation energy to break the bond if the energy were developed uniformly throughout the material. However, such a high energy density is not expected, since it would break all of the bonds. If the molecular weight of a polymer is 1 800 000 and its bulk density is 0.9 g/cm^3, its bulk molar volume is 2 000 000 cm^3. The actual volume of a single polymer molecule is then 3.3×10^{-18} cm^3. The volume in which the molecule can be enveloped is greater than this even in bulk, and much greater in a 10 percent solution. Therefore, the energy necessary to break one bond can be developed mechanically in a volume less than one-tenth of the smallest possible volume which could be pervaded by a single polymer molecule. It should be pointed out that the utilization of energy is not uniform over this volume but is greatest by far in those regions of the volume actually occupied by the molecule. The largest possible segment of a polymer chain which could occupy this necessary volume would have about 10 000 carbon atoms. If the energy developed in this volume could be concentrated into a single bond, the activation energy necessary to break the bond would be supplied. Such concentration does not seem unfeasible, if one considers the nature of the flow of viscoelastic materials. Whereas in the flow of ordinary liquids the shearing energy applied is dissipated as heat immediately when it is applied, in the flow of viscoelastic materials the finite times required for individual inter- and intramolecular movements result in some of the applied shearing energy being temporarily stored as potential energy in the molecules for the finite periods of time required for the occurrence of the individual molecular movements. This temporarily stored potential energy is distributed nonuniformly over the various chemical bonds present in a macroscopic sample of material, since the individual molecular movements required for macroscopic flow are of varied nature throughout the sample. Since the distribution of this temporarily stored potential energy is nonuniform it will contribute a high energy to some of the bonds present, and these bonds may find it possible to undergo rupture. One simple way by which a given bond could be raised to a high-energy level would be by the combination of tensions originating in several polymer molecular chains into one through a multiple entanglement of the molecular chains. While these comparisons are not very definite, they do indicate at least a possible energy balance.

[9] Reference 8, p. 15.

VI. SUMMARY

Moderately concentrated solutions of polyisobutylenes having molecular weights above about 500 000 show pronounced viscosity decreases when sheared at a nominal rate of shear of 65 000 sec^{-1} at 37.8°C. Similar solutions of lower molecular weight polyisobutylenes show much smaller viscosity decreases.

Apparent recoveries of viscosity on resting after shearing are indicated to be spurious, instrumental effects.

The viscosity decreases are shown to correspond to molecular chain degradations by the fact that the intrinsic viscosity of a solution measured after shearing is less than the value measured before shearing.

The energy required for the molecular chain ruptures can be supplied by a reasonable concentration of the energy dissipated in the shearing process.

ACKNOWLEDGMENT

The authors' appreciation is expressed to the Enjay Company, Inc. who furnished the polymers investigated.

Nonlinear Semiconductor Resistors

F. A. SCHWERTZ AND J. J. MAZENKO
Mellon Institute, Pittsburgh 13, Pennsylvania*
(Received March 9, 1953)

A phenomenological theory for the nonlinear voltage-current characteristic curve displayed by a granular aggregate of silicon carbide is presented. The theory, which should apply to granular semiconductors other than silicon carbide, is based on a very simple model of the aggregate and on the assumption that the essential resistance is located at the grain-grain contacts, the impedance of the bulk material being considered negligibly small. The current i is related to the voltage V by the equation

$$i=kAP^{n/m}d^{n-2}V^n/t^n,$$

where A is the cross-sectional area of the aggregate, t the thickness, P the applied pressure, and d the average particle diameter. The constants k, n, and m are structure-sensitive, that is, they depend on the physical-chemical nature of the particles forming the aggregate. Factors influencing the values of these constants include impurity concentration in the semiconductor, elastic constants, and particle shape. The validity of this equation is supported by careful measurements made on granular aggregates of silicon carbide.

INTRODUCTION

A GRANULAR aggregate of silicon carbide displays a nonlinear voltage-current characteristic curve which is frequently described by the empirical equation

$$i=k'V^n. \tag{1}$$

In this expression i is the current flowing through the aggregate, V is the applied voltage, and k' and n are constants. For a given type of silicon carbide, the value of the constant k' depends on the dimensions of the cell containing the aggregate, the average particle diameter, and the applied pressure. The constant n appears to be independent of these parameters, and its magnitude generally lies between five and seven.

It is shown here that the constant k' may be represented by the expression

$$k'=kAP^{n/m}d^{n-2}/t^n, \tag{2}$$

where A is the cross-sectional area of the cell containing the granular aggregate, t the thickness of the cell, that is, the distance between the electrodes applied to the aggregate, P the applied pressure, and d the average particle diameter. The value of the constant m, for silicon carbide, is about three and one-half. Equation (1) may therefore be written in the form

$$i=kP^{n/m}Ad^{n-2}V^n/t^n, \tag{3}$$

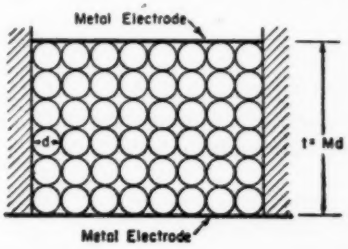


FIG. 1. Simplified model of granular semiconductor resistor.

where k is a constant whose value is characteristic of the particular type of silicon carbide employed in the measurements. Equation (3) is general in nature and should apply to semiconductors other than silicon carbide, due allowance being made for variations in the structure-sensitive constants k, m, and n.[1,2]

PHENOMENOLOGICAL THEORY

Equation (3) may be derived with the aid of the empirical equation (1) and the following simplifying assumptions:

(a) The granular aggregate may be regarded as a parallel set of N noninteracting grain pillars, each pillar containing M individual, identical grains (Fig. 1).

(b) The essential resistance lies at the grain-grain and grain-metal (electrode) interfaces, the resistance of the bulk material being relatively negligible.[3]

Where the thickness of the granular aggregate t is much larger than the particle diameter d, the resistance of the $(M-1)$ grain-grain contacts will predominate and the resistance of the metal-grain contacts will be relatively negligible. If, however, as a slightly more accurate approximation, it is assumed that the two metal-grain contacts are not negligible but equivalent to a grain-grain contact, there will be M effective grain-grain contacts. Put in another way, each grain acts as a nonlinear resistor obeying the law

$$i_j=k_jv_j^n. \tag{4}$$

If there are M of these in series, the total voltage drop V across the column will be given by

$$V=\sum_{j=1}^{M} v_j=Mv_j. \tag{5}$$

Combining these two equations gives

$$i_j=k_j[V/M]^n. \tag{6}$$

* Multiple Fellowship on Computer Components, sustained by the Electronics Directorate of the U. S. Air Force Cambridge Research Center, Research Development Command, under Contract AF 19/122/-376.

[1] John Bardeen, Phys. Rev. **71**, 717 (1947).
[2] Taylor, Odell, and Fan, Phys. Rev. **88**, 867 (1952).
[3] Else Holm, J. Appl. Phys. **23**, 509 (1952).

TABLE I. Particle size classification of silicon carbide grain.

Passed through sieve number	Retained on sieve number	Mean particle diameter in inches
80	100	0.0065
120	140	0.0045
170	200	0.0032
230	270	0.00225

If now there are N identical pillars of grain in parallel, the total current i through the granular column will be

$$i=Ni_j=k_jN[V/M]^n. \tag{7}$$

Since M and N are related to the average particle diameter d, the cell thickness t, and the cell area A by the equations

$$Md=t, \tag{8}$$

and

$$N\doteq 4A/\pi d^2, \tag{9}$$

Eq. (7) yields

$$i=4k_jAd^{n-2}V^n/\pi t^n. \tag{10}$$

We have yet to consider the influence of pressure on the impedance of the grain-grain contacts. To do this we first relate the pressure applied to the granular aggregate P, to the average pressure p actually existing at the grain-grain contacts. For the case of two spheres of equal diameter d in contact with each other, or for a point of diameter d in contact with an infinite plane, p is given by the expression[4]

$$p=\text{const}[fE^2/d^2]^{\frac{1}{3}}, \tag{11}$$

where f is the total force at the grain-grain contact and E is the modulus of elasticity in compression. If now F is the total force or weight applied to the granular aggregate,

$$f=\frac{F}{N}=\frac{\pi d^2F}{4A}=\frac{\pi d^2P}{4}. \tag{12}$$

Placing this value for f into Eq. (11) yields

$$p=\text{const}P^{\frac{1}{3}}, \tag{13}$$

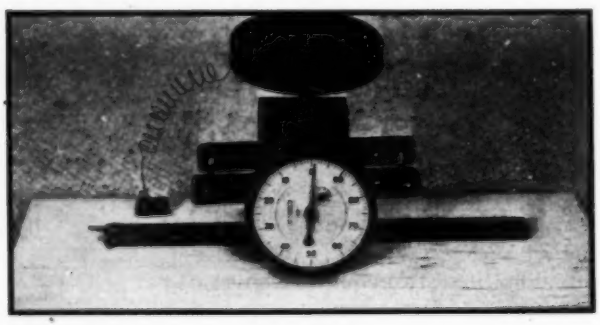


FIG. 2. Experimental setup for impedance measurements.

[4] S. Timoshenko, *Theory of Elasticity* (McGraw-Hill Book Company, Inc., 1934), first edition, p. 343.

a result which shows that the average pressure at the grain-grain contacts is independent of the particle diameter.

We now invoke the experimental fact that

$$\left[\frac{\partial \log V}{\partial \log P}\right]_i=-\frac{1}{m}, \tag{14}$$

where the experimental values of m are in the neighborhood of three to four. It follows that

$$V=\text{const}/P^{1/m} \tag{15}$$

at constant current. Braun and Busch[5] have also examined the pressure-dependence of the grain-grain contact resistance, and they assign a value of three to m on the basis of their measurements. A value of three implies, in virtue of Eq. (13), that the voltage required to produce a constant current varies inversely as the actual average pressure at the grain-grain contacts, or that

$$v_i=\text{const}/p. \tag{16}$$

With the aid of Eq. (4) we can therefore state that

$$i_j=k_j[\text{const}/p]^n \tag{17}$$

TABLE II. n values for various particle diameters.

Average particle diameter, inches	n values
0.0065	5.66
0.0045	5.60
0.0032	5.51
0.00225	5.62
	Average 5.60

at constant current. Since the choice of i_j is arbitrary it follows that

$$k_j=k_1p^n=k_2P^{n/3}, \tag{18}$$

where k_1 and k_2 are constants. If m is not assigned the value 3 but is left arbitrary, Eq. (18) is replaced by

$$k_j=k_2P^{n/m}. \tag{19}$$

Placing this value of k_j into Eq. (10) yields

$$i=k\frac{P^{n/m}Ad^{n-2}V^n}{t^n}, \tag{3}$$

where the factor $(4/\pi)$ has been absorbed into the constant k. A more compact formula is obtained by using the current density

$$j=i/A, \tag{20}$$

and applied field

$$E=V/t, \tag{21}$$

for in this case

$$j=kP^{n/m}d^{n-2}E^n. \tag{22}$$

[5] A. Braun and G. Busch, Helv. Phys. Acta 15, 574 (1942).

CURRENT I, MILLIAMPERES

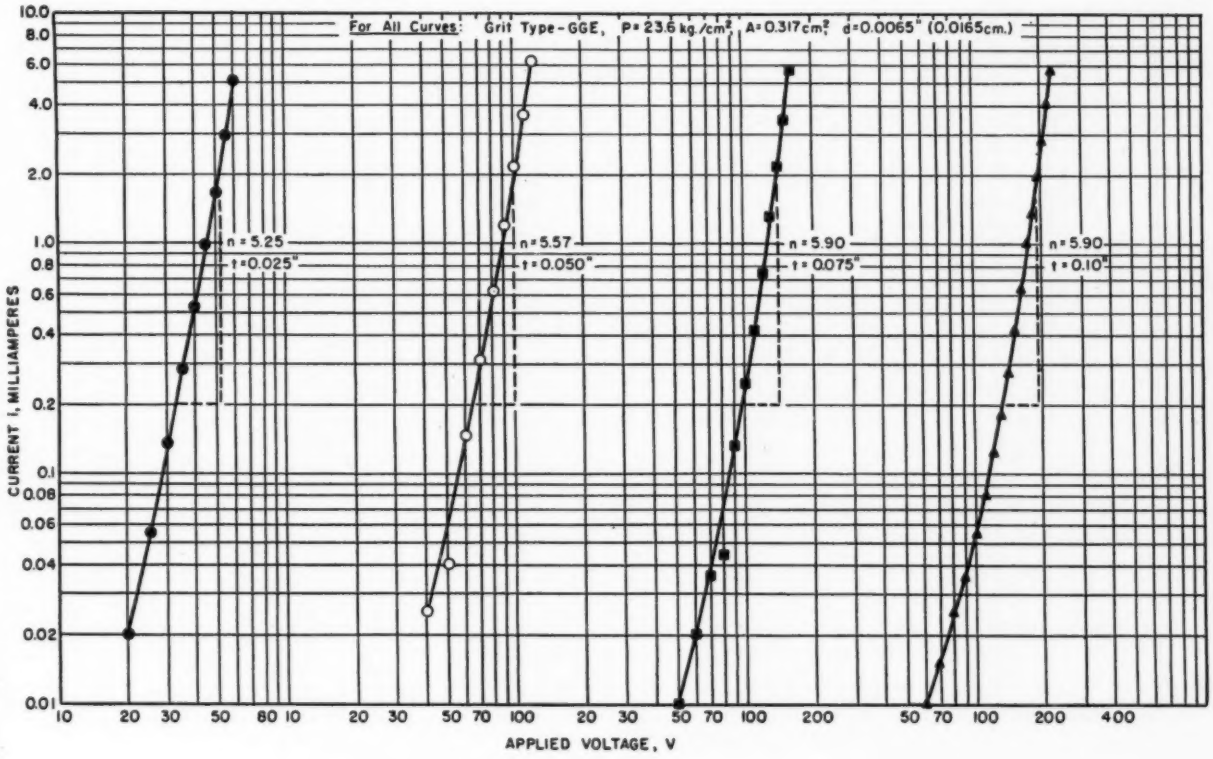


FIG. 3. Voltage-current characteristic curves.

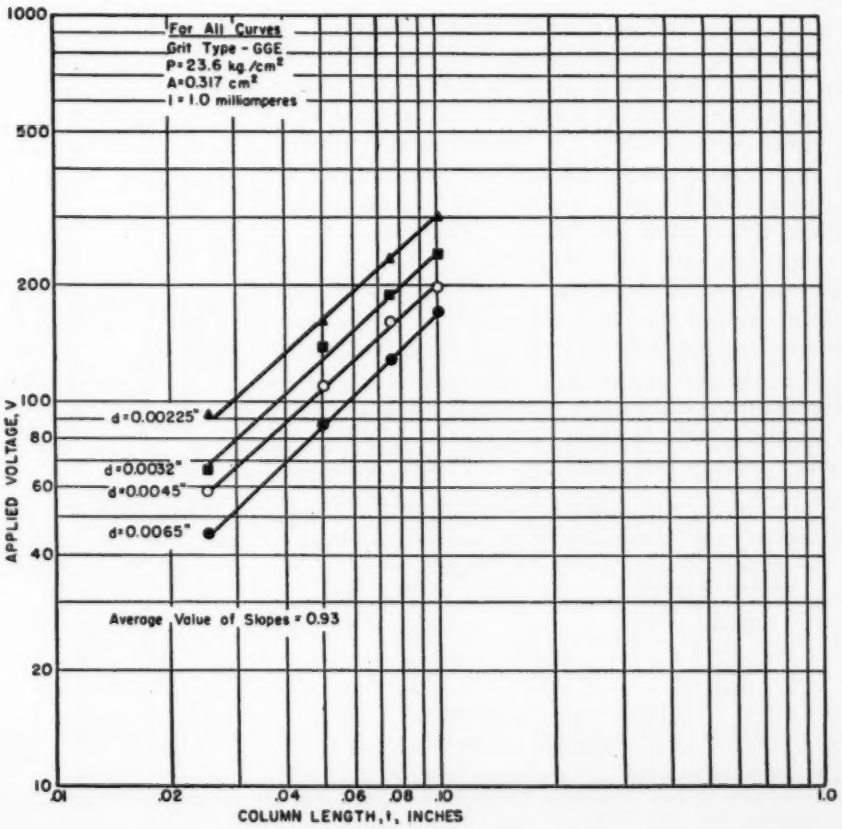


FIG. 4. Voltage at constant current as a function of column length.

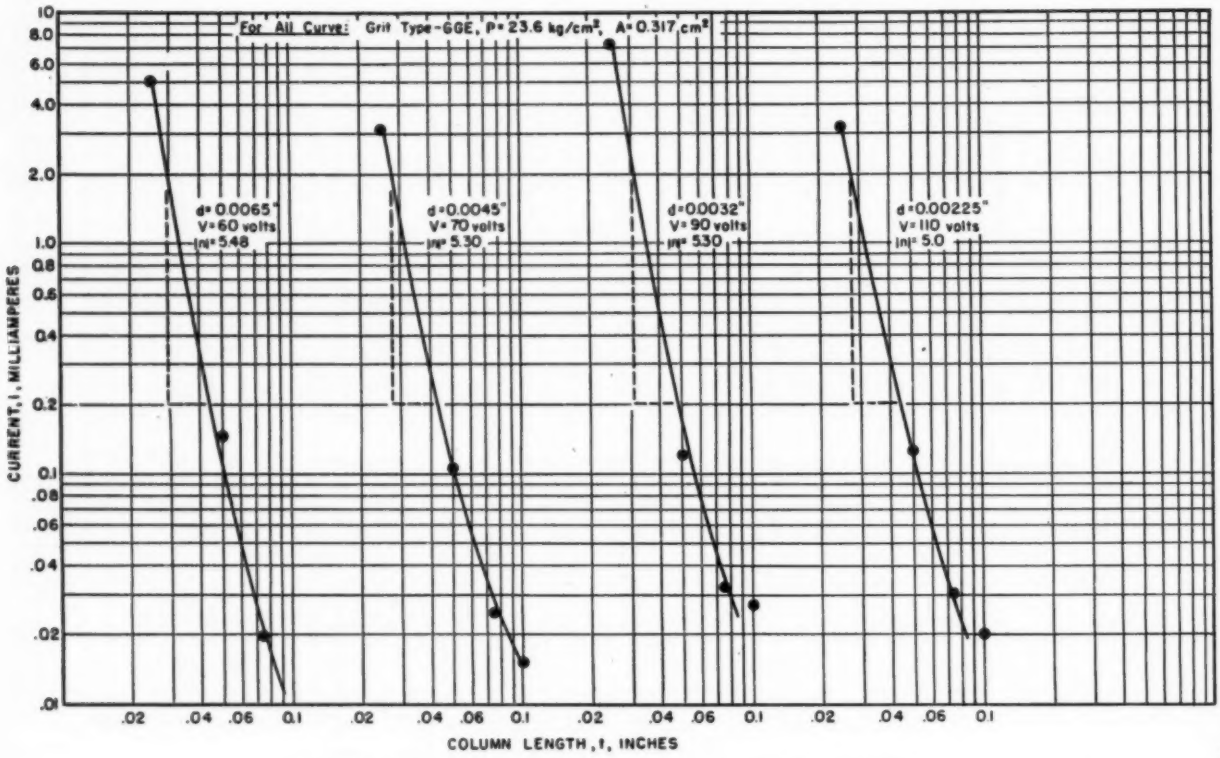

For All Curve: Grit Type-GGE, P=23.6 kg/cm², A=0.317 cm²
d=0.0065" V=60 volts n=5.48
d=0.0045" V=70 volts n=5.30
d=0.0032" V=90 volts n=5.30
d=0.00225" V=110 volts n=5.0
CURRENT, i, MILLIAMPERES
COLUMN LENGTH, l, INCHES

FIG. 5. Current at constant voltage as a function of column length.

MATERIALS AND APPARATUS

Most of the experimental measurements were made on a black-type silicon carbide grain designated by the maker† as 120GGE electrical grade. This material was fractionated with the aid of U. S. Standard fine sieves and the fractions listed in the following table were retained for experimental purposes. The figures for the mean particle diameter entered in the third column were obtained by averaging the standard values given for the sieve openings.

In order to obtain the voltage-current characteristics of the granular silicon carbide, the grain was placed into a $\frac{1}{4}$ in. hole drilled in a Micarta block (Fig. 2). The block itself was fastened to a steel plate which served as the bottom electrode; the top electrode consisted of a rod of cold-rolled steel which slid freely into the hole. This rod was made movable in order to apply pressure to the granular column conveniently. A "dial indicator" graduated in 0.001-in. divisions was used to measure the length of the column. With this assembly the effects of

TABLE III. n values at fixed voltages.

Average particle diameter d, inches	Slope, n, of i vs l curves
0.0065	5.48
0.0045	5.30
0.0032	5.30
0.00225	5.00
	Average 5.27

† The silicon carbide grain used in these experiments was kindly supplied by Mr. G. Van Nimwegen of The Carborundum Company, Niagara Falls, New York.

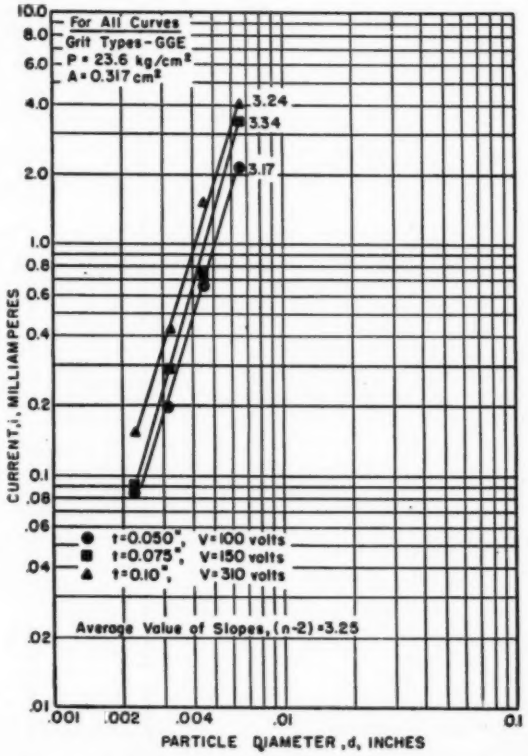


FIG. 6. Current at constant voltage as a function of particle diameter.

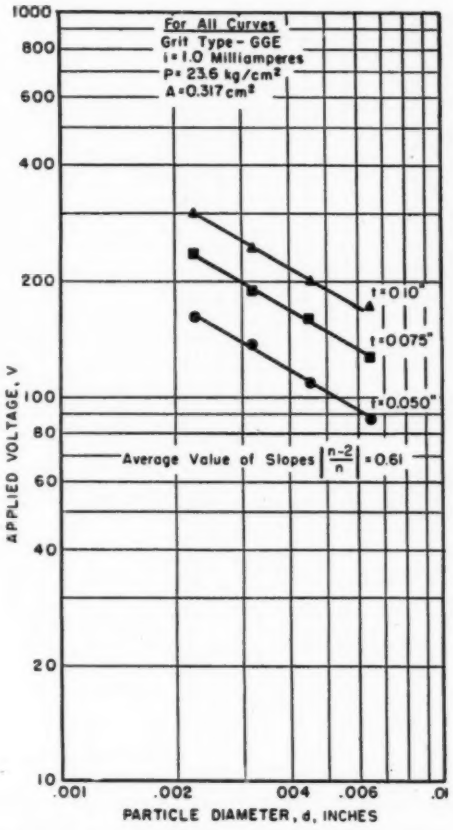


FIG. 7. Voltage at constant current as a function of particle diameter.

length, pressure, and particle size on the impedance of the granular column were very readily determined.

EXPERIMENTAL PROCEDURE

The experimental evidence on which Eq. (3) is based was gathered by keeping either the applied voltage or the current constant while varying one of the several experimental parameters occurring on the right-hand side of Eq. (3). The results so obtained were represented graphically on log-log plots. In order to keep matters straight, the fixed values of the nonvaried parameters were recorded on each graph. In addition, the type of silicon carbide grain employed in the measurements was also noted. All the measurements were made at room temperature.

The pressure was applied to the top electrode of the assembly shown in Fig. 2. An aspirator bottle filled with mercury was placed on the horizontal "shelf"

TABLE IV. $(n-2)$ values at fixed voltages.

Column thickness t, inches	Slope, $(n-2)$, of i vs d curves
0.050	3.17
0.075	3.34
0.10	3.24
	Average 3.25

TABLE V. $(n-2)/n$ values at fixed current.

Column thickness, inches	Slope, $(n-2)/n$, of V vs d curves
0.050	0.615
0.075	0.608
0.100	0.610
	Average 0.61

attached to this electrode, and the pressure was regulated by allowing the mercury to run out, in controlled amounts, through a stopcock located near the base of the bottle. Attempts to vary the pressure with the aid of a group of metal weights were unsuccessful because the slight jar experienced by the top electrode, upon adding or removing one of the weights, was sufficient to change the nature of the grain-electrode and grain-grain contacts within the granular mass. After the pressure was applied, and before the voltage-current readings were taken, the silicon carbide was "settled-out" by rapping the Micarta housing with a series of sharp blows. Where these precautions were taken, results of a reproducible nature were obtained.

As mentioned above, most of the measurements were made on black silicon carbide grain. This variety of grain is known to behave as a p type semiconductor, whereas the green variety behaves as n type.[6] As a

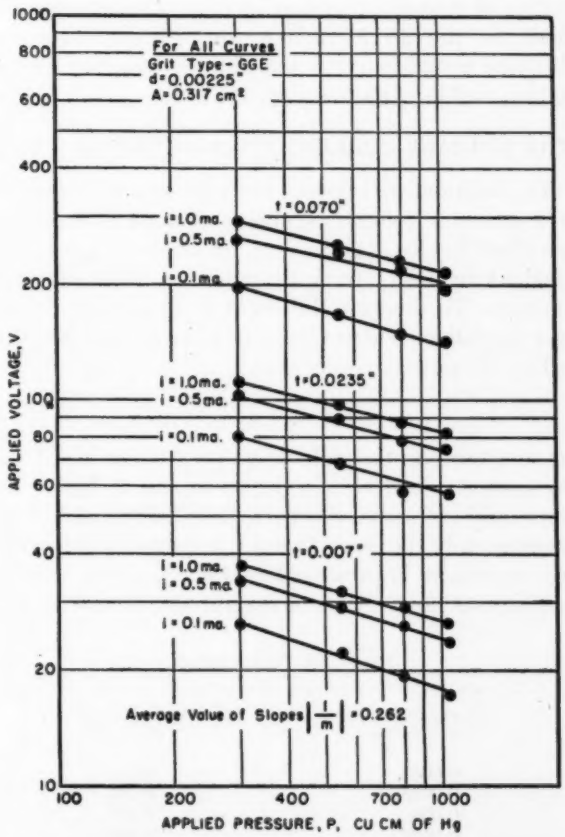


FIG. 8. Voltage at constant current as a function of pressure.

[6] G. Busch and G. Labhart, Helv. Phys. Acta 19, 463 (1946).

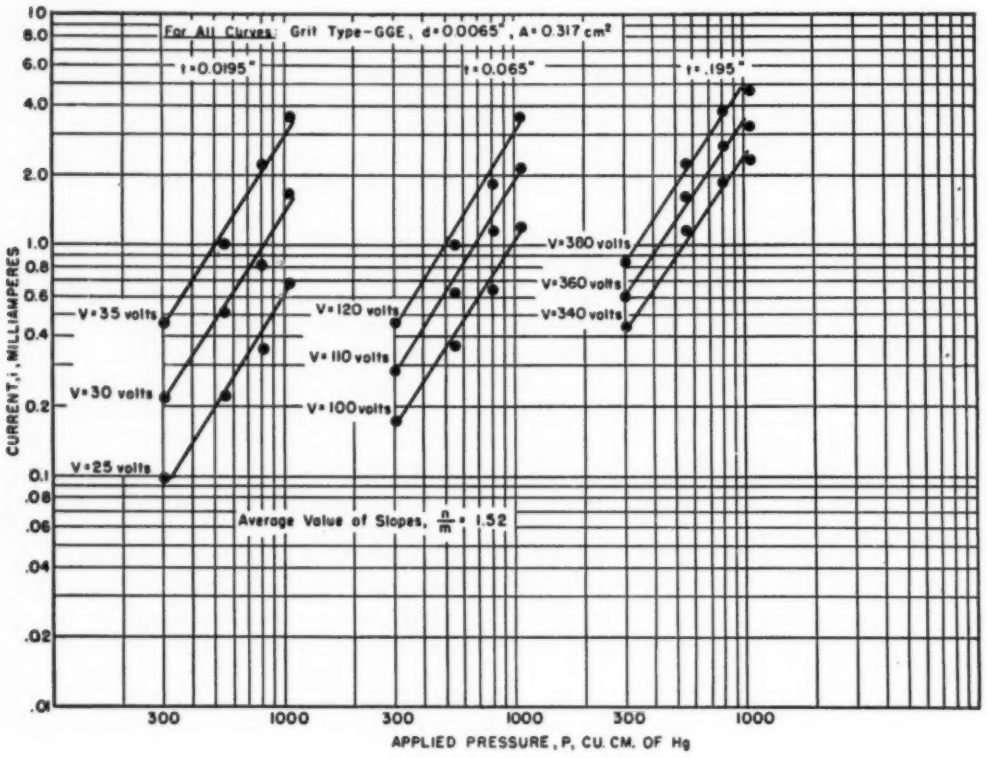


FIG. 9. Current at constant voltage as a function of pressure.

matter of interest, therefore, a few measurements were made on the green variety. In addition, data were obtained on grain types differing in impurity concentration and in particle shape.

THE VOLTAGE-CURRENT CHARACTERISTIC CURVE

The relationship between the voltage and the current in a granular column of silicon carbide is accurately described by Eq. (1) only at reasonably high current densities or field strengths where n is substantially constant. To describe the voltage-current relationship over the full current range, a linear term must be added to Eq. (1) so that

$$i=k_1'V+k_2'V^n, \qquad (23)$$

the first term on the right-hand predominating at very low currents (and voltages), the second at higher currents (and voltages). For currents of practical interest, however, only the second term is required. Evidence for this statement is presented in Fig. 3, where currents ranging over four orders of magnitude, from 0.01 to 10 milliamperes, are plotted against applied voltage on logarithmic scales. Over this whole range the voltage-current curves are practically straight lines. Consequently in all the experiments described here, interest is confined to the straight portion of the characteristic curve where Eq. (1) adequately describes the voltage-current relationship.

The index of nonlinearity n was determined for the curves in Fig. 3 by measuring the average value of the slope of each in the current range from 0.2 to 2.0 milliamperes. The average value of the slopes so obtained was 5.66. As indicated in Fig. 3, this average applied to silicon carbide particles of average diameter 0.0065 in. and to column lengths ranging from 0.025 in. to 0.10 in. Similar figures, drawn for the other particle diameters, produced the same value for n within experimental error. The results are summarized in Table II.

EFFECT OF COLUMN LENGTH

According to Eq. (3), the voltage required to produce a constant current is proportional to the column length

TABLE VI. $(1/m)$ values.

Average particle size, d, inches	Average values of $(1/m)$
0.0065	0.281
0.0045	0.293
0.0032	0.276
0.00225	0.268
	0.280 Average

TABLE VII. (n/m) values.

Average particle size, inches	Average values of (n/m)
0.0065	1.52
0.0045	1.59
0.0032	1.62
0.00225	1.53
	1.57 Average

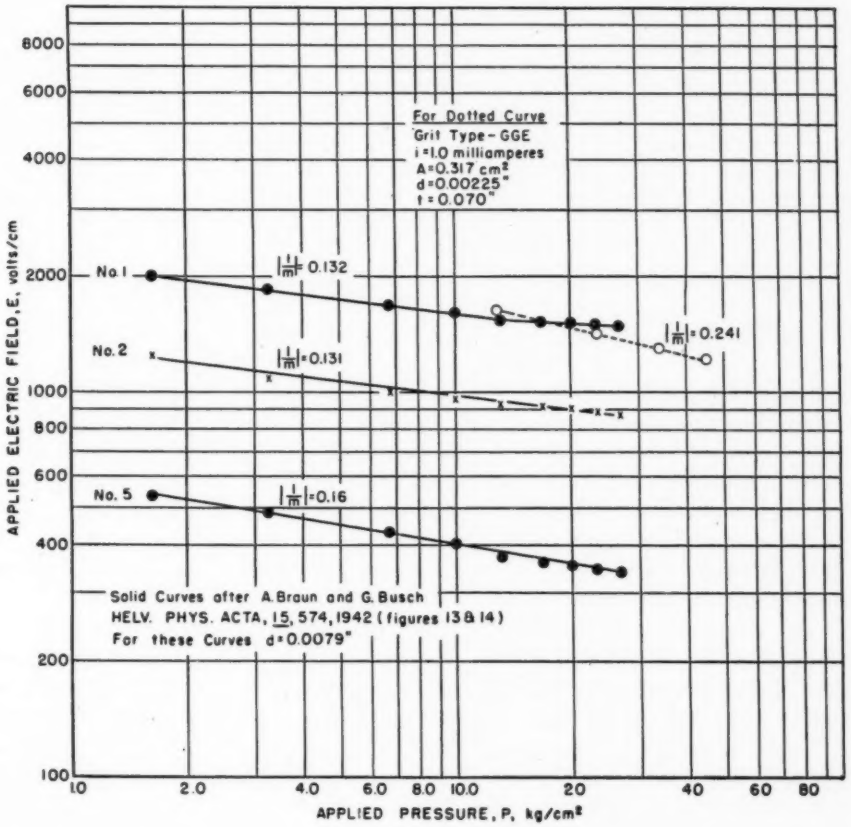

For Dotted Curve
Grit Type - GGE
i = 1.0 milliamperes
A = 0.317 cm²
d = 0.00225"
t = 0.070"
No. 1
|1/m| = 0.132
|1/m| = 0.241
No. 2
|1/m| = 0.131
No. 5
|1/m| = 0.16
Solid Curves after A. Braun and G. Busch
HELV. PHYS. ACTA, 15, 574, 1942 (figures 13 & 14)
For these Curves d = 0.0079"
APPLIED ELECTRIC FIELD, E, volts/cm
APPLIED PRESSURE, P, kg/cm²

FIG. 10. Electric field at constant current as a function of pressure.

t. A logarithmic plot of the applied voltage *vs* the length should therefore produce a straight line of slope unity. Lines of this sort are presented in Fig. 4 for each of the four particle sizes used in the experiments. The average value for the slopes of these lines is 0.93, or unity within experimental error.

The current produced at constant voltage should vary inversely as the nth power of the column length. Accordingly a logarithmic plot of the current *vs* length should produce a straight line of slope n. Four lines of this sort are shown in Fig. 5, one for each particle size employed in the experiments. The magnitudes of the slopes of these lines are summarized in Table III. The average value of 5.27 for n compares favorably with the average value obtained directly from the i *vs* V curves, namely 5.60.

EFFECT OF PARTICLE DIAMETER

According to Eq. (3) the current produced at constant voltage should vary directly as the $(n-2)$th power of the average diameter. A logarithmic plot of the current *vs* average particle diameter should therefore produce a straight line of slope slightly greater than three. Three lines of this sort are presented in Fig. 6. In Table IV the slopes obtained from these curves are seen to yield an average value for $(n-2)$ of slightly over 3, as predicted.

The voltage required to produce a constant current should vary inversely as the $(n-2)/n$ power of the average particle diameter. A logarithmic plot of the voltage *vs* average particle diameter should therefore produce a straight line of slope approximately equal to 0.6. Three such lines are presented in Fig. 7. In Table V the slopes of these are summarized and shown to yield an average value $(n-2)/n=0.61$.

EFFECT OF PRESSURE

It was found very difficult to ascertain, in a reproducible way, the effects of pressure on the voltage-

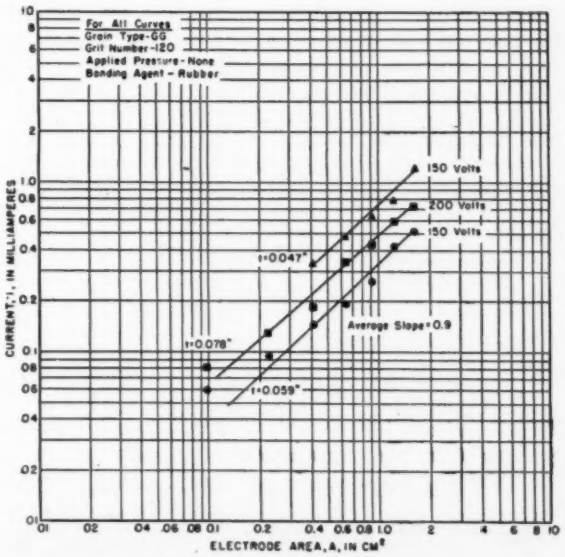


FIG. 11. Current as a function of area at constant voltage.

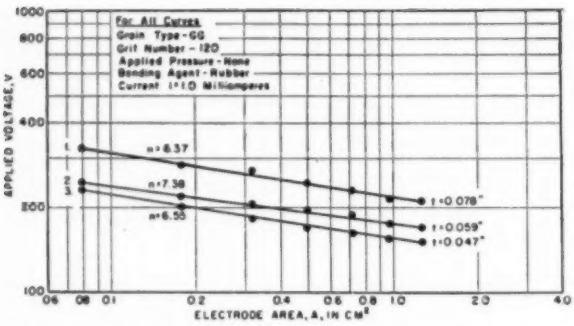


FIG. 12. Voltage as a function of area at constant current.

current characteristics curve of silicon carbide. Satisfactory results were obtained only where the pressure was *decreased* stepwise. Where the pressure was increased in a stepwise fashion, erratic results were always obtained. Apparently the silicon carbide granules shifted position under the influence of increasing pressure so that, in effect, one was looking at new resistor for each experimentally determined point.

The procedure employed in the experiments was to determine the voltage required to produce a given current as the pressure was varied, then to determine the current at constant voltage. At constant current, the voltage should vary inversely as the $(1/m)$th power of the applied pressure. Results of this sort are displayed in Fig. 8 where the voltage *vs* pressure data are represented in a log-log plot for several column thicknesses and for several values of the current. The data refer to particles which passed through a No. 230 sieve and were retained on a No. 270 sieve. As noted above, their average diameter was 0.00225 inch. Since the electrode diameter for all of these measurements was 0.25 inch ($A=0.317$ cm²) the pressure was not plotted directly along the horizontal axis; instead the number of cubic centimeters of mercury in the aspirator bottle was used for this purpose, due regard being taken to the weight of the bottle and the electrode system. The weight of these was equivalent to 50 cc of mercury. The pressure values in cc of mercury may be converted to kg/cm² by employing the factor 0.043.

The values of the slopes of the lines shown on Fig. 8 are very sensitive to the manner in which the experimental points are interpreted, and significance therefore should be attached only to their average value. For Fig. 8 this average value is $1/m=0.262$ or m approximately equal to four. Similar figures were constructed for each particle size; the average values of the slopes for these are summarized in Table VI. These values lead to a grand average of 0.280 for $(1/m)$, and hence $m=3.57$.

Equation (3) requires that the current produced at constant voltage be proportional to the (n/m)th power of the applied pressure. Curves showing this type of behavior are presented in Fig. 9 where the current has been plotted logarithmically against the applied pressure. Here again only the average value of the slopes is significant. This average is tabulated below for the nine curves on Fig. 9 along with the similar values obtained for the other particle sizes. As seen from Table VII, the grand average for the slopes is 1.57. Using this value for (n/m) and the value obtained above for m, namely 3.57, gives 5.60 for n.

As mentioned earlier, Braun and Busch[5] find that the voltage required to produce a constant current varies inversely as the cube root of the applied pressure, that is, they select for m the value three. In their paper the data on which this selection is based are displayed by plotting the applied voltage (actually the applied field) *vs* the reciprocal of the cube root of the applied pressure. This is an *insensitive* way of representing the results since all the data for pressures between unity and infinity are compressed into the domain between zero and unity along the pressure axis. If, however, the data are displayed according to the scheme of Fig. 7 where the logarithm of the applied voltage (actually the applied field) is plotted against the logarithm of the applied pressure the results shown in Fig. 10 are obtained. The full lines represent the data of Braun and Busch; the slopes of these curves correspond to a pressure exponent of approximately $-(1/7)$ rather than $-(1/3)$. The dotted line represents the topmost curve of Fig. 8, but it has been reproduced in Fig. 10 using the same *field* and

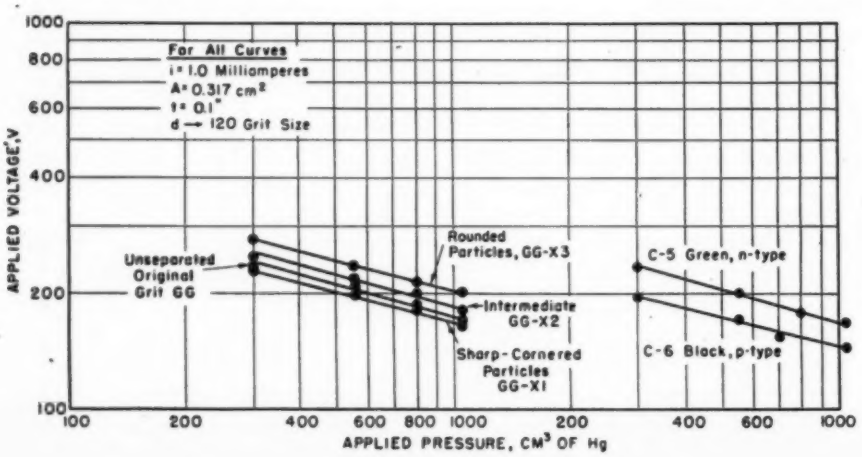


FIG. 13. Voltage *vs* pressure curves for various silicon carbide types.

TABLE VIII. Silicon carbide grain-grit No. 120.

Sample No.	Grit type	Color	Remarks
1	GGE	Black	a
2	C-5	Green	b
3	C-6	Black	c
4	GG	Black	d
5	GGX1	Black	e
6	GGX2	Black	e
7	GGX3	Black	e

[a] Special grain for electrical applications.
[b] The C-5 silicon carbide is prepared from all virgin materials with no added refurnaced products. Because of the purity of the raw materials, the resulting product comes out with a light green color.
[c] The C-6 silicon carbide contains a higher percentage of impurities, as well as refurnaced materials, and as a result has a black color.
[d] The GG silicon carbide is first quality black silicon carbide which has been mechanically treated to produce grains of high bulk packing density.
[e] Samples 5, 6, and 7 were obtained by separating the 120 GG grain according to particle shape into three fractions on a sloping and vibrating tilt-table. Microscopic observations showed that the GG-X1 particles were flat and splintery in shape, whereas the GG-X3 particles were very roughly elliptical in shape and were relatively rounded-off at the corners. GG-X2 was intermediate in character.

applied pressure units as those employed by the above-mentioned authors.

Values of the pressure exponent about equal to those found by Braun and Busch were obtained during the course of the present work in those situations where measurements were made under conditions of increasing pressure. In these cases, however, the results were always erratic, and the curves often showed kinks or knees, which presumably indicated particle movements within the granular mass.

It appears reasonable to conclude that the best value for m is about 3.6.

EFFECT OF ELECTRODE AREA

Equation (3) requires that the current produced at constant voltage be proportional to the electrode area. Although this requirement appears almost obvious, its experimental verification is quite difficult. This is so because the determination of each experimental point on an i *vs* A curve involves a wholly new experimental set-up, requiring, among other things, a new set of electrodes and a new sample of grit. The experimental errors, accumulated in all the necessary changes, made the acquisition of reliable data quite difficult. It was therefore decided to obviate these complications by employing flat rubber-bonded sheets of granular silicon carbide as the test specimens. Electrodes of diameters varying from $\frac{1}{8}$ in. to $\frac{1}{2}$ in., increasing in $\frac{1}{16}$ in. steps, were placed on the sheet surfaces by spraying conducting silver paint through suitably formed stencils. The current produced at constant voltage and the voltage produced at constant current were then determined. The data obtained at constant voltage were displayed by plotting the current *vs* electrode area on a log-log basis. For this type of representation a straight line of slope unity should be obtained. Three such lines are shown in Fig. 11 for three different sheet thicknesses.

TABLE IX. n and m values for various silicon carbide types.

Sample No.	Silicon carbide type	n	m
1	GGE	5.6	3.6
2	C-5	5.6	3.6
3	C-6	5.6	4.0
4	GG	6.8	3.6
5	GG-X1	6.8	3.6
6	GG-X2	6.8	3.6
7	GG-X3	6.8	3.9

The average value for the slopes of these three lines is 0.9.

By plotting the voltage against the area on a log-log basis, lines with slope of magnitude $(1/n)$ should be obtained. Three such lines are displayed in Fig. 12, one for each rubber-bonded sheet. The n values calculated from the slopes are shown in the figure. Their average value is 6.8. As will be shown in the next section, this is the value to be expected for the GG grit employed in the rubber-bonded sheets.

EFFECT OF SILICON CARBIDE TYPE

The constants k, n, and m are structure-sensitive and depend on such factors as impurity concentration in the semiconductor and on particle shape. Several samples of No. 120 grit which differed with respect to these properties were therefore examined experimentally. Data relative to these samples are summarized in Table VIII.

The index of nonlinearity n for each of the samples listed in Table VIII was obtained directly from the voltage-current characteristic curves. Values of m were obtained by logarithmically plotting voltage *vs* pressure curves at constant current. These latter curves are shown in Fig. 13. Unfortunately, this figure contains no data relative to the electrical grade grit GGE, since this material was not available when the data in this figure were collected. The n and m values are tabulated in Table IX. No relative values of k are presented in this table because no information was available concerning the particle size distribution within the various samples, and, although this consideration does not affect the m and n values, it does influence the apparent k values.

The m values in Table IX appear to be about the same within experimental error, indicating that the influence of pressure on the voltage-current characteristic curve is identical for all grit types. The high value for n possessed by the GG grit-type is undoubtedly real, and its value is apparently independent of particle shape. On the other hand, it can be stated on the basis of the data presented in the left-hand side of Fig. 13 that the k values do depend on particle shape, for these results show that a lower voltage is required to produce a given current in an aggregation of sharp-cornered particles than in an assembly of rounded-off particles, other things being equal. This phenomenon was previously observed by Braun and Busch,[5] and they at-

tributed it to the high local electric fields developed at the points and corners of the splintery particles.

CONCLUSIONS

The phenomenological theory for the voltage-current characteristics of granular silicon carbide appears to be adequately supported by the data presented here. The theory should apply equally well to granular aggregates of other semiconductors in the pressure ranges where the granules behave as elastic bodies, due allowance being made for variations in the structure-sensitive constants k, n, and m.

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 8 AUGUST, 1953

Filters for Detection of Small Radar Signals in Clutter

HARRY URKOWITZ
Philco Corporation Research Division, Philadelphia, Pennsylvania
(Received April 6, 1953)*

Radar clutter is distinguished from thermal noise by being caused by random reflection of transmitted electromagnetic energy. This paper considers what may be done by means of linear filtering to improve the detection of small radar signals in the presence of this clutter. Defining the signal-to-clutter ratio as the ratio of the peak signal to the rms value of the clutter, the optimum linear filter is derived for enhancing this ratio. The optimum filter has a frequency characteristic given by the conjugate of the voltage spectrum of the transmitted pulse divided by the power spectrum of the clutter. The synthesis of the optimum filter and various approximations to it are considered. The influence of these filters upon the signal-to-thermal-noise ratio is also investigated. Finally, the influence of the second detector upon the signal-to-clutter or signal-to-noise ratio is considered.

1. THE OPTIMUM FILTER

CLUTTER in radar systems is distinguished from fluctuation noise by being caused by the reflection of transmitted electromagnetic energy. Radar return from rain drops, radar reflection from ground or sea surface as seen by a moving airplane, and reflection from metallized strips formerly used as a radar jamming measure, all constitute forms of clutter as it is considered here. In many systems, clutter is the major source of unwanted interference. We shall assume that clutter is the only form of interference to be considered.

We shall see that the second detector suppresses small signals in the presence of clutter, the smaller the signal, the greater the relative suppression. This is the reason for wanting to improve the signal-to-clutter ratio before the second detector. This paper therefore considers the synthesis of a linear filter with a frequency characteristic such that the output, measured at the input to the second detector, will have the maximum signal-to-clutter ratio which is measured by the ratio of the peak signal to the root-mean-square value of the clutter. Actually, we shall use the square of this ratio. After finding the desired frequency characteristic, and discussing its synthesis we consider various approximations to the optimum. We also consider the noise behavior of these filters.

A very much simplified diagram of a pulsed radar system is shown in Fig. 1. The reflected radio energy at high frequency (RF) is heterodyned with the local oscillator output in the mixer, whose frequency differs from the RF by an amount equal to the intermediate frequency (IF). The IF amplifier is followed by the second detector, which usually consists of a rectifier followed by a low-pass filter. The output of the second detector may be considered to be equal to the envelope of the input if we use a linear, half-wave rectifier, and equal to the square of the envelope if we use a full-wave quadratic rectifier. We call these the linear and square law detectors, respectively. The system may be considered to be completely linear up to the second detector, and it is assumed that no frequency restriction occurs until the IF section is reached. This latter assumption is not very stringent, since we may consider all RF effects at IF. Because the system is linear, we can talk of the clutter and the signal separately up to the second detector. The question to be answered can be stated as follows: What should the frequency characteristic of the IF section be to maximize the ratio of the peak signal voltage (or current) to the mean square value of the clutter voltage (or current), at the input to the second detector?

This question has already been answered for a general case. If the interference is noise with a flat power spectrum, the optimum filter has a frequency characteristic given by the conjugate of the voltage spectrum of the received pulse. This result was derived by D. O. North,[1] who used a method based on the calculus of variations, and independently by Van Vleck and

* The original, unrevised manuscript was received December 11, 1952.

[1] D. O. North, Technical Report PtR-6C, RCA Laboratories, Princeton, New Jersey, June 25, 1943.

Middleton,[2] who used a method based on Schwarz's inequality.

If the power spectrum of the noise or interference is not flat, a generalization of the previous result is necessary. This has been accomplished by T. S. George,[3] who used a method similar to North's, and independently by B. M. Dwork,[4] who applied Schwarz's inequality. Their result is that, except for a constant time delay, the optimum filter is given by the conjugate of the voltage spectrum of the received pulse divided by the power spectrum of the interference.

Let $A(\omega)$ = frequency characteristic of filter, $E(\omega)$ = voltage spectrum of received signal pulse, $E^*(\omega)$ = conjugate of $E(\omega)$, $F(\omega)$ = power spectrum of the interference.

Then

$$A(\omega)=k_1E^*(\omega)/F(\omega), \tag{1}$$

where k_1 is a proportionality constant. In order to solve our problem, we must find $F(\omega)$, the power spectrum of the clutter. This has already been found. Lawson and Uhlenbeck[5] state that the power spectrum of the received clutter is the same as that of the transmitted pulse. T. S. George[3] derives this result by treating the transmitted pulse as a scanning aperture. The power spectrum of the transmitted pulse is, of course, proportional to $|E(\omega)|^2=E(\omega)E^*(\omega)$. Then

$$A(\omega)=k_1/E(\omega). \tag{2}$$

For the sake of definiteness, let the received signal be a rectangular pulse of RF, with amplitude unity, and duration δ.

$$E(\omega)=\int_{-\delta/2}^{\delta/2}\cos\omega_o t\,\exp\{-j\omega t\}dt$$

$$=\frac{\delta}{2}\left[\frac{\sin(\omega-\omega_o)\delta/2}{(\omega-\omega_o)\delta/2}+\frac{\sin(\omega+\omega_o)\delta/2}{(\omega+\omega_o)\delta/2}\right]. \tag{3}$$

The input signal-to-clutter ratio is

$$\left(\frac{S}{C}\right)_1=\frac{1}{\frac{1}{2\pi}\int_{-\infty}^{\infty}|E(\omega)|^2d\omega}. \tag{4}$$

Any constants of proportionality will cancel later when we compare this with the output signal-to-clutter ratio. Substituting Eq. (3) into Eq. (4), one obtains

$$(S/C)_1=2/\delta. \tag{5}$$

The filter to which the signal and clutter are applied has a frequency characteristic given by Eq. (2), but in order to be more realistic, we suppose that the filter cuts off beyond a certain frequency on either side of the center frequency. This is then, a truncated version of the filter of Eq. (2). Nevertheless, it is labeled "optimum" because we recognize that it is more realistic to consider a band-limited filter than one whose desirable characteristics extend to infinity in both directions. Thus,

$$A_{\text{opt}}(\omega)=k_1/E(\omega), \quad \text{for} \quad \omega_o-\omega_c<|\omega|<\omega_o+\omega_c \tag{6}$$
$$=0, \quad \text{otherwise.}$$

Multiplying Eq. (6) by Eq. (3) and transforming, the maximum value of signal is found to be

$$e_2(t)_{\max}=2k_1\omega_c/\pi. \tag{7}$$

The clutter output is

$$C_2=\frac{1}{2\pi}\int_{-\omega_o-\omega_c}^{-\omega_o+\omega_c}k_1^2d\omega+\frac{1}{2\pi}\int_{\omega_o-\omega_c}^{\omega_o+\omega_c}k_1^2d\omega$$

$$=2\omega_ck_1^2/\pi. \tag{8}$$

Then, the output signal-to-clutter ratio is

$$\left(\frac{S}{C}\right)_2=\frac{[e_2(t)_{\max}]^2}{C_2}=\frac{2\omega_c}{\pi}=4f_c. \tag{9}$$

Then, the improvement provided by the filter is the ratio of Eq. (9) to Eq. (5):

$$R_{\text{opt}}=\frac{(S/C)_2}{(S/C)_1}=2f_c\delta. \tag{10}$$

Equation (10) states that the improvement in signal-to-clutter ratio provided by the optimum filter is equal to the product of pulse length and band width. It should not be concluded that increasing the pulse length will improve the visibility of the signal, because the input clutter power is proportional to the pulse length. All that one can do is to increase the band width. The actual magnitude of the input signal-to-clutter ratio will depend on many factors, such as target reflectivity of the scatterers which cause the clutter, etc.

2. COMPARISON WITH FLAT FILTER

Before considering the synthesis of the optimum filter, it is desirable to compare the clutter behavior of a

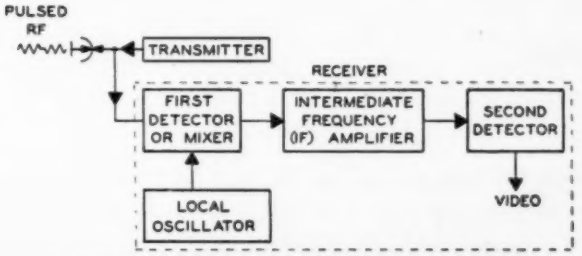


FIG. 1. Simplified diagram of a pulsed radar system.

[2] J. Van Vleck and D. Middleton, J. Appl. Phys. **17**, 940–971 (1946).

[3] T. S. George, Philco Corp. Research Division, Report No. 159, Feb. 1950; Inst. Elec. Eng. (London), Monograph No. 22 (1952).

[4] B. M. Dwork, Proc. Inst. Radio Engrs. **38**, 771–774 (1950).

[5] J. L. Lawson and G. E. Uhlenbeck, *Threshold Signals*, MIT Radiation Laboratory, Series No. 24 (McGraw-Hill Book Company, Inc., New York, 1950), p. 297.

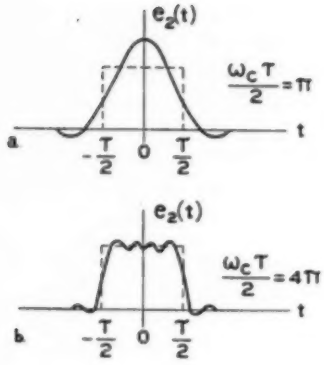


FIG. 2. The function $[Si(\omega_c[t+\frac{1}{2}\delta]) - Si(\omega_c[t-\frac{1}{2}\delta])]$. All τ's in the figure should be δ's.

flat filter. Of course, an actual filter only approximates the sharp cutoff assumed here. Also, if one wishes to avoid overshoots, the skirts of the filter are made to slope gradually. To find the output signal-to-clutter ratio of the flat filter, we may use the band-pass, low-pass analogy, and work with the envelopes. Then the frequency characteristic of the flat filter is:

$$\begin{aligned} |A_F(\omega)| &= 1, \quad |\omega| < \omega_c, \\ &= 0, \quad |\omega| > \omega_c. \end{aligned} \tag{11}$$

Furthermore, we assume linear phase (constant time delay for all frequencies). Now, we have

$$E(\omega)=\delta\frac{\sin\omega\delta/2}{\omega\delta/2}, \tag{12}$$

$$F(\omega)=\delta^2\left(\frac{\sin\omega\delta/2}{\omega\delta/2}\right)^2. \tag{13}$$

We want the ratio

$$R_F=\frac{\left[\frac{1}{2\pi}\int_{-\omega_c}^{\omega_c}A_F(\omega)E(\omega)\exp\{j\omega t\}d\omega\right]^2}{\frac{1}{2\pi}\int_{-\omega_c}^{\omega_c}|A_F(\omega)|^2F(\omega)d\omega} \tag{14}$$

where the numerator is evaluated at its maximum value. The denominator (clutter output) is easily evaluated. Substituting Eqs. (11) and (13) into the denominator of Eq. (14), there is obtained

$$C_2=\delta I\left(\frac{\omega_c\delta}{2}\right) \tag{15}$$

where

$$I(x)=\frac{2}{\pi}\int_0^x\left(\frac{\sin t}{t}\right)^2dt.$$

Also, we have, using Eq. (11) and Eq. (12),

$$e_2(t)=\frac{1}{\pi}[Si(\omega_c[t+\delta/2])-Si(\omega_c[t-\delta/2])], \tag{16}$$

where $Si(x)$ is the sine integral function.[6] The function inside the brackets of Eq. (16) is plotted for two different values of $\omega\delta/2$ in Fig. 2. It can be seen that the maximum value of $e_2(t)$ does not necessarily occur at $t=0$. For small values of $e_2(t)$, say $\omega_c\delta/2\leqslant\pi$, the peak value of $e_2(t)$ occurs at $t=0$:

$$e_2(t)_{\max}=\frac{2}{\pi}Si\left(\frac{\omega_c\delta}{2}\right). \tag{17}$$

As $\omega_c\delta/2$ increases, $Si(\omega_c\delta/2)$ increases to a value 1.85 at $\omega_c\delta=2\pi$ and then oscillates around the value $\pi/2$ with decreasing amplitude.[6] This oscillation represents overshoot or ringing. As mentioned before, this overshoot will not occur if the filter does not cut off sharply. For $\omega_c\delta>2\pi$, the second term in the brackets of Eq. (16) is very close to $\pi/2$, when the first is near its maximum value, and vice versa. Thus, for $\omega_c\delta>2\pi$,

$$\begin{aligned} e_2(t)_{\max} &\approx \frac{1}{\pi}\left[Si(\omega_c[t+\delta/2])_{\max}+\frac{\pi}{2}\right] \\ &= \frac{1}{\pi}\left[Si(\pi)+\frac{\pi}{2}\right] \\ &= \frac{3.4}{\pi}. \end{aligned} \tag{18}$$

This appears to be independent of $(\omega_c\delta/2)$, which means that once overshoot occurs, the percentage of overshoot will remain approximately constant with increasing band width. This is the occurrence of the Gibbs phe-

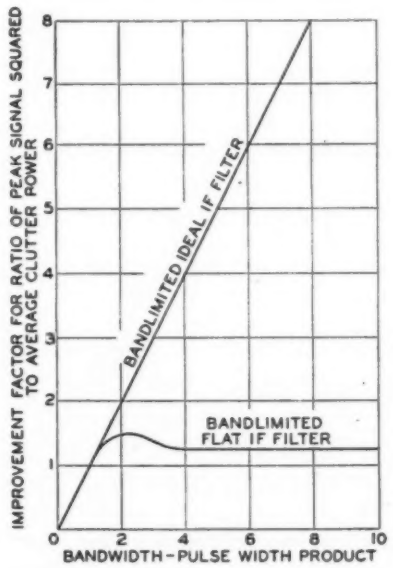


FIG. 3. Comparison of improvement of signal-to-clutter ratio for band-limited, ideal IF filter and band-limited, flat filter.

[6] E. Jahnke and F. Emde, *Tables of Functions* (Dover Publications, New York, 1945), p. 1.

nomenon in the Fourier integral.[7] Then using Eqs. (15), (17), and (18), it is found that

$$\left(\frac{S}{C}\right)_2=\frac{4Si^2(\omega_c\delta/2)}{\pi^2\delta I(\omega_c\delta/2)}, \tag{19}$$

for $\omega_c\delta\leqslant 2\pi$, and

$$\left(\frac{S}{C}\right)_2=\frac{11.7}{\pi^2\delta I(\omega_c\delta/2)} \tag{20}$$

for $\omega_c\delta>2\pi$. By the usual technique it is found that

$$\left(\frac{S}{C}\right)_1=\frac{1}{\delta}, \tag{21}$$

$$\therefore\quad R_F=\frac{(S/C)_2}{(S/C)_1}=\frac{11.7}{\pi^2 I(\omega_c\delta/2)} \tag{22}$$

for $\omega_c\delta>2\pi$, and

$$R_F=\frac{4Si^2(\omega_c\delta/2)}{\pi^2 I(\omega_c\delta/2)} \tag{23}$$

for $\omega_c\delta\leqslant 2\pi$.

R_F and R_{opt} are plotted in Fig. 3. To recapitulate, R_{opt} is the improvement factor for the bandlimited optimum filter and R_F is the improvement factor of the flat filter of equal band width. Note that the flat filter has practically no effect on the signal-to-clutter ratio beyond a certain band-width-pulse length product.

3. SYNTHESIS OF OPTIMUM FILTER

A first step in the synthesis is drawing the pole-zero pattern in the complex frequency plane. We must know the filter characteristic as a function of the complex frequency, $s=\alpha+j\omega$. This is found by writing the characteristic as a function of $j\omega$, that is, as a Fourier transform, and then substituting $s=\alpha+j\omega$ for $j\omega$. The optimum filter characteristic for a rectangular pulse is found by using Eq. (6) and Eq. (3), and is, except for a constant factor, with s for $j\omega$,

$$A_{\text{opt}}(s)=\frac{s-j\omega_o}{\sinh(s-j\omega_o)\frac{\delta}{2}}. \tag{24}$$

This function has no zeros, but has poles where

$$s=j[\omega_o\pm(2\pi n/\delta)],\quad n\neq 0 \tag{25}$$

in which n is an integer. Of course, these are poles with infinite Q; they are on the $j\omega$ axis. The gain function is shown in Fig. 4.

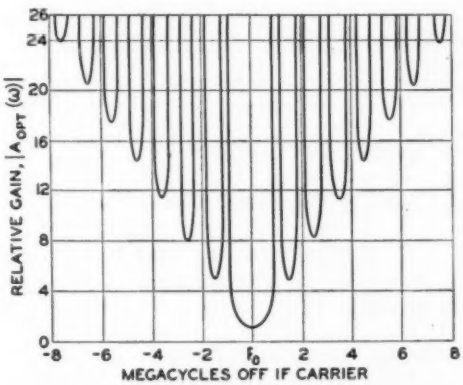


FIG. 4. Frequency response of optimum anticlutter filter for a transmitted pulse of one μsec duration.

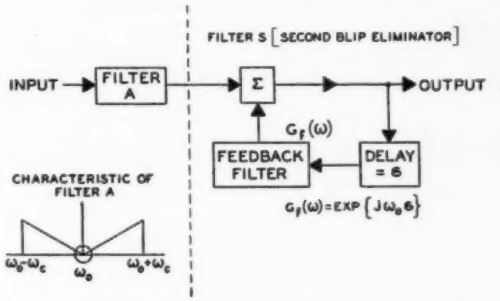


FIG. 5. Synthesis of optimum filter.

One way to achieve the optimum characteristic is to cascade two networks, one of which has a single zero of infinite Q at the center of the IF passband, the other having poles of infinite Q, one at the center of the passband and the others separated from it and each other by an angular frequency, $(2\pi/\delta)$. This is illustrated in Fig. 5. Filter A has a single infinite Q zero at the center of the IF passband. The characteristic of filter A is shown with a sharp cutoff at $\omega_o\pm\omega_c$. As will be shown later, the output of filter A is two IF pulses, each of a duration less than that of the input pulse; the second pulse has its phase reversed with respect to the first. The first of the shortened pulses goes through filter S unmolested, but the second gets canceled by the first after it is delayed. Thus, filter S is a "second blip eliminator." The gain function of the second blip eliminator is, except for a constant factor,

$$G_S(s)=\frac{1}{1-\exp[-(s-j\omega_o)\delta/2]}. \tag{26}$$

We find that the gain function of the second blip eliminator has poles of infinite Q at the center frequency, and at angular frequencies, $\omega_o\pm(2\pi n/\delta)$. Therefore, since the center pole is canceled by the zero of filter A, the pole pattern of filter A cascaded with filter S is the same as that of the optimum filter.

The "filter with zero," which is our designation for filter A, is interesting in its own right. We shall show that filter A by itself provides an improvement in signal-to-clutter ratio which is half that provided by

[7] S. Goldman, *Frequency Analysis, Modulation, and Noise* (McGraw-Hill Book Company, Inc., New York, 1948), p. 80.

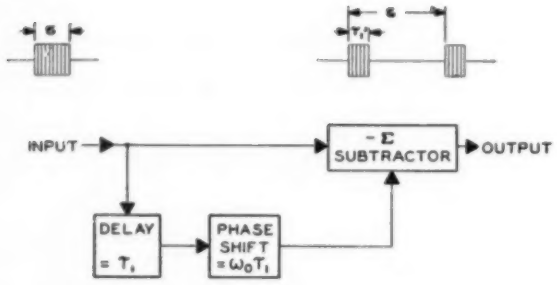


FIG. 6. Pulse length aperture correction network.

the optimum filter of equal band width, that is, three decibels less in signal-to-clutter ratio. The assumption of linear phase is not justified in this case, because of the zero on the axis at the center of the passband. The phase characteristic is discontinuous at ω_o and has a jump equal to π radians. This fact and reference to Fig. 5 leads us to assume that

$$A_A(\omega)=k_1|\omega-\omega_o|\exp\left\{-j\frac{\pi}{2}\right\}, \text{ for } \omega_o-\omega_c<\omega<\omega_o$$

$$=k_1|\omega-\omega_o|\exp\left\{j\frac{\pi}{2}\right\}, \text{ for } \omega_o<\omega<\omega_o+\omega_c. \quad (27)$$

If we carry out the computation for the envelope of the output of filter A, by multiplying $E(\omega)$ by $A_A(\omega)$ and taking the inverse transform, it is found that

$$e_2(t)=\frac{k_1}{\pi}\left[\frac{\sin\omega_c(t+\frac{1}{2}\delta)}{t+\frac{1}{2}\delta}-\frac{\sin\omega_c(t-\frac{1}{2}\delta)}{t-\frac{1}{2}\delta}\right]. \quad (28)$$

Thus the output is two $(\sin x/x)$ IF pulses separated by the pulse duration δ. Each of these pulses is of shorter duration than the input pulse, the shortening depending upon the band width. For ω_c large enough, $e_2(t)$ has a maximum at $t=-\frac{1}{2}\delta$:

$$e_2(t)_{\max}\approx k_1(2f_c)=k_1\times\text{band width}. \quad (29)$$

The clutter power output is found by integrating $F(\omega)|A_A(\omega)|^2$ where $F(\omega)$ is the clutter power spectrum given by Eq. (13) and $A_A(\omega)$ is given by the low-frequency analog of Eq. (27). The result is

$$C_2=\frac{4k_1^2}{\pi}\left[\frac{\omega_c}{2}+\frac{\sin\omega_c\delta}{2\delta}\right]. \quad (30)$$

As an approximation, valid for $\omega_c\delta>\pi$, we write

$$C_2\approx 2k_1^2\omega_c/\pi. \quad (31)$$

Then the output signal-to-clutter ratio is obtained by dividing the square of Eq. (29) by Eq. (31). Dividing this ratio by Eq. (21) we obtain

$$R_A=f_c\delta=\tfrac{1}{2}\text{ (band width }\times\text{ pulse duration)}. \quad (32)$$

This confirms our previous statement that filter A is three decibels worse in signal-to-clutter ratio than the optimum filter of equal band width. The remaining improvement comes from the second blip eliminator.

4. PULSE LENGTH APERTURE CORRECTION

An approximation to filter A, the "filter with zero," may be obtained by a delay and subtraction network, as illustrated in Fig. 6. We call this filter a "pulse length aperture corrector." The gain function of this filter is proportional to

$$|A_B(\omega)|=|\exp\{-j(\omega-\omega_o)\tau_1\}-1|, \quad (33)$$

with a similar expression in $\omega+\omega_o$ for the image (negative) frequencies. The phase shift is again discontinuous. We have, in the neighborhood of $+\omega_o$,

$$\text{phase shift}=-\pi/2, \text{ for } \omega<\omega_o, \ =\pi/2, \text{ for } \omega>\omega_o. \quad (34)$$

In the neighborhood of $-\omega_o$ we have,

$$\text{phase shift}=-\pi/2, \text{ for } \omega<-\omega_o, \ =\pi/2, \text{ for } \omega>-\omega_o,$$

since the phase shift is an odd function of frequency. This filter has zeros at ω_o and at angular frequencies separated from ω_o and each other by $2\pi/\tau_1$. We choose the band width to be $2/\tau_1$, so that $\omega_c=2\pi/\tau_1$. We assume that the gain is zero outside $\omega_o\pm\omega_c$. A similar statement applies to the image frequencies. To find the output envelope, we apply the technique used before, taking image frequencies into account. The result is

$$e_2(t)=\frac{2\delta}{\pi}\left[\int_0^{\omega_c}\frac{\sin(\frac{1}{2}\omega\delta)}{(\frac{1}{2}\omega\delta)}\cos\omega t\,d\omega - \int_0^{\omega_c}\frac{\sin(\frac{1}{2}\omega\delta)}{(\frac{1}{2}\omega\delta)}\cos\omega(t-\tau_1)d\omega\right]. \quad (35)$$

For effective operation, $\tau_1\ll\delta$, or $\omega_c\gg 2\pi/\delta$. Thus, the two terms of Eq. (35) represent two blips, each of a duration τ_1, approximately, and the second displaced from the first by a time equal to δ. Each of the integrals has been evaluated before. We find that

$$e_2(t)_{\max}=3.4/\pi. \quad (36)$$

The power gain function is given by

$$|A_B(\omega)|^2=|(1-\exp\{-j(\omega-\omega_o)\tau_1\})|^2 =2(1-\cos[\omega-\omega_o]\tau_1). \quad (37)$$

As before, we write

$$C_2=\frac{2\delta^2}{2\pi}\int_{-\omega_c}^{\omega_c}\left[\frac{\sin\frac{1}{2}\omega\delta}{\frac{1}{2}\omega\delta}\right]^2(1-\cos\omega\tau_1)d\omega. \quad (38)$$

Now since $\tau_1\ll\delta$, and $\omega_c=(2\pi/\tau_1)$, we see that outside the interval $(-\omega_c,\omega_c)$, the integrand is very small.

Therefore,[8]

$$C_2=\frac{2\delta^2}{\pi}\int_0^\infty\left[\frac{\sin(\frac{1}{2}\omega\delta)}{(\frac{1}{2}\omega\delta)}\right]^2(1-\cos\omega\tau_1)d\omega=2\tau_1. \quad (39)$$

From Eq. (36),

$$e_2(t)_{\max}\approx 1. \quad (40)$$

The excess is overshoot. Therefore

$$(S/C)_2=1/2\tau_1=f_c/2, \quad (41)$$

and

$$R_B=f_c\delta/2=\delta/2\tau_1. \quad (42)$$

Of course, the units of Eqs. (39) and (41) are incorrect, but these come out right when the ratio, R, is found. Equation (42) tells us that the improvement in signal-to-clutter ratio of the pulse length aperture corrector is just 3 decibels less than the IF filter with zero of equal band width.

5. NOISE BEHAVIOR OF ANTICLUTTER FILTERS

Each of the filters considered above is compared to a flat filter of equal band width. We assume that the noise input has a value of k_2 volts (rms) per unit angular band width. For the flat filter, whose frequency characteristic is given by Eq. (11), we find the noise power output by multiplying the total band width by k_2.

$$N_2=4k_2\omega_c. \quad (43)$$

This takes into account the image pass band centered at $-\omega_o$. Then, for the flat filter, using Eq. (18)

$$\left(\frac{S}{N}\right)_F=\frac{2.9}{\pi^2k_2\omega_c}. \quad (44)$$

For the optimum filter, we use

$$N_2=\frac{k_2}{2\pi}\int_{\omega_o-\omega_c}^{\omega_o+\omega_c}|A(\omega)|^2d\omega$$

$$=\frac{k_2}{\pi\delta}\int_0^{\omega_c}\left[\frac{(\frac{1}{2}\omega\delta)}{\sin(\frac{1}{2}\omega\delta)}\right]^2d\omega. \quad (45)$$

This integral grows very large as ω_c increases. In fact, for $\omega_c>(2\pi/\delta)$,

$$N_2=\infty.$$

Then

$$\left(\frac{S}{N}\right)_{\text{opt}}=0, \quad \text{for} \quad \omega_c>(2\pi/\delta). \quad (46)$$

For filter A, the IF filter with zero, we use similar methods, and we find that

$$\left(\frac{S}{N}\right)_A=\frac{3}{4k_2\pi^2\omega_c}. \quad (47)$$

Comparing this with Eq. (45) for the flat filter, we find that

$$\frac{(S/N)_A}{(S/N)_F}\approx\tfrac{1}{4}\text{ or }-6\text{ decibels}. \quad (48)$$

In a similar manner, we find, for the pulse length aperture corrector, that

$$\frac{(S/N)_B}{(S/N)_F}=\tfrac{1}{2}\text{ or }-3\text{ decibels}. \quad (49)$$

The presence of anticlutter filters in the IF section is particularly helpful when one considers the effect of the second detector in suppressing small signals. We shall show that the suppressing effect of the second detector is particularly bad for very small signals. The phenomenon of small signal suppression in noise has been treated by others.[9] Here we wish to show that a similar thing occurs in clutter, and we wish to derive quantitative results for our definitions of signal-to-clutter ratio.

6. ACTION OF THE SECOND DETECTOR IN THE PRESENCE OF CLUTTER

The radar return which we have labeled "clutter" may be considered to be the sum of the returns from many small scatterers located in a certain volume of space (or area of the earth's surface). The dimensions of this volume are determined by the pulse length and the antenna beam width. The radial extent of this volume is equal to $(c\delta/2)$, where c is the velocity of light. Randomness is introduced by assuming that the phases of the individual returns are random and uniformly distributed over the interval $(0, 2\pi)$. Then, the returned current may be written as

$$I(t)=\sum_{n=1}^{N}\{C_n\cos(\omega_o t-\phi_n)\}, \quad (50)$$

where C_n is the amplitude of the nth component, ω_o is the angular frequency of the returned energy translated to intermediate frequency, and ϕ_n is the random phase angle. Equation (50) may be written as

$$I(t)=\sum_{n=1}^{N}\{C_n\cos\phi_n\cos\omega_o t+C_n\sin\phi_n\sin\omega_o t\}$$

$$=I_c\cos\omega_o t+I_s\sin\omega_o t, \quad (51)$$

where

$$I_c=\sum_{n=1}^{N}C_n\cos\phi_n,$$

$$I_s=\sum_{n=1}^{N}C_n\sin\phi_n. \quad (52)$$

Let us assume that N is very large, that the contribution

[8] Bierens de Haan, *Nouveau Tables d'Integrales Definies* (G. E. Stechert and Company, New York, 1939), table no. 157, formula no. 5.

[9] D. Middleton, Proc. Inst. Radio Engrs. **36**, 1467–1477 (1948).

of any one component is small, and that the scatterers are statistically independent. Then the central limit theorem[10] of statistics allows us to state that I_s and I_c are two normally distributed variables. They are also independent, since $\langle I_s I_c \rangle_{Av}=0$. The envelope of $I(t)$ is defined by

$$R(t)=(I_c^2+I_s^2)^{\frac{1}{2}}. \tag{53}$$

So far, the statistical description of clutter is the same as that for noise given by Rice.[11] We shall use Rice's results calculated for noise when only the first probability density functions of $I(t)$ are involved, that is in the calculation of the moments.

Let us assume that a linear detector is one which produces an output voltage (or current) which is equal to the IF envelope. To define the signal in the video, we use the idea of Jordan.[12] During most of the sweep cycle nothing but clutter (or noise) is supplied to the detector which builds up an average voltage, $\bar{R}_c$, across the detector load. Then, for the pulse duration, say one microsecond or a fraction thereof, a carrier of amplitude P, due to the target, is applied to the detector. For the duration of the pulse, the average voltage is increased to some value, $\bar{R}_s$. (We have assumed a rectangular pulse.) This situation is illustrated in Fig. 7. The following formula is obtained from Rice.[11]

$$R_T=\bar{R}_s-\bar{R}_c=\left(\frac{\psi_o\pi}{2}\right)^{\frac{1}{2}}[{}_1F_1(-\tfrac{1}{2};1;-x)-1], \tag{54}$$

where ψ_o is the mean square value of the clutter at the input to the second detector, ${}_1F_1$ is the confluent hypergeometric function,[13,14] and x is the signal to clutter-*power* ratio in the IF:

$$x=P^2/2\psi_o. \tag{55}$$

According to our definition of signal-to-clutter ratio, we have

$$\text{Signal-to-clutter ratio}=P^2/\psi_o=2x. \tag{56}$$

For $x\ll 1$,

$$R_T\approx\frac{P^2}{4}\left(\frac{\pi}{2\psi_o}\right)^{\frac{1}{2}}. \tag{57}$$

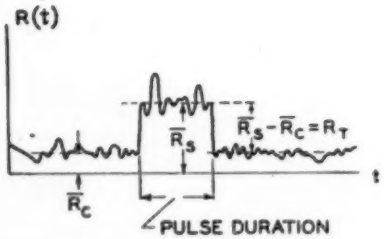


FIG. 7. Typical video trace of signal and clutter.

As far as the visibility of the target return is concerned, it is the presence of clutter on either side of the target return which obscures visibility. Since the signal is considered to be the average departure from the clutter, we consider that the interference arises from the fluctuation around $\bar{R}_c$, the average of the clutter envelope. What we desire, then, is the mean square value around $\bar{R}_c$. This is the dispersion. Using Rice's results, we have, as the dispersion,

$$\langle (R_c-\bar{R}_c)^2\rangle_{Av}=\bar{R}_c^2-(\bar{R}_c)^2$$

$$=\psi_o\left(.2-\frac{\pi}{2}\right)=0.43\psi_o. \tag{58}$$

The video signal-to-clutter ratio is the square of Eq. (54) divided by Eq. (58):

$$\left(\frac{S}{C}\right)_{\text{vid}}=\frac{R_T^2}{\langle (R_c-\bar{R}_c)^2\rangle_{Av}}$$

$$=\frac{\pi}{4-\pi}[{}_1F_1(-\tfrac{1}{2};1;-x)-1]^2. \tag{59}$$

The effect of the second detector upon signal-to-clutter ratio is determined by dividing Eq. (59) by $2x$, the input signal-to-clutter ratio. Then we have

$$F_{\text{vid}}=\left(\frac{S}{C}\right)_{\text{vid}}\Big/\left(\frac{S}{C}\right)_{\text{IF}}$$

$$=\frac{\pi[{}_1F_1(-\tfrac{1}{2};1;-x)-1]^2}{2(4-\pi)x}. \tag{60}$$

For $x\ll 1$, we divide the square of Eq. (57) by $2x$ to obtain

$$F_{\text{vid}}(\text{small } x)\approx\pi x/8(4-\pi). \tag{61}$$

Equation (61) shows clearly the reduction of signal-to-clutter ratio for small signals. It is of interest to note the behavior of the second detector for large signal to clutter ratio. For very large x, we use an asymptotic expansion for the confluent hypergeometric function. For x of large magnitude,[14]

$${}_1F_1(-\tfrac{1}{2};1;-x)\approx 2(x/\pi)^{\frac{1}{2}}. \tag{62}$$

Then,

$$\lim_{x\to\infty}F_{\text{vid}}=\frac{2}{4-\pi}=2.33, \tag{63}$$

or 3.68 decibels. F_{vid} as a function of $2x$ is plotted in Fig. 8. A curve such as Fig. 8 enables one to determine the true worth of IF anticlutter devices. For example, suppose that the signal-to-clutter ratio in the IF is -4 decibels. If the filter does not change this ratio, the output of the second detector would be 10 decibels less than this, or -14 decibels. Now, suppose that we provide an IF filter which increases the signal-to-clutter

[10] M. G. Kendall, *Advanced Theory of Statistics* (C. Griffin and Company, London, 1948), Vol. 1, fourth edition, p. 180.
[11] S. O. Rice, Bell System Tech. J. **24**, 75 ff. (1945).
[12] W. H. Jordan, Mass. Inst. Technol. Radiation Lab. Report 61–23, July, 1942.
[13] Reference 6, p. 275.
[14] Reference 10, Appendix 4-B.

ratio by six decibels, making the ratio +2 decibels at the input to the second detector. The signal-to-clutter ratio would be decreased 5 decibels by the second detector, making the output signal-to-clutter ratio equal to −3 decibels. This is a gain of 11 decibels over the previous case. The desirability of IF improvement is clear.

7. DISCUSSION

Filters A and B discussed above provide good improvement in signal-to-clutter-ratio, and do not deteriorate the noise performance too much. Filter A increases the signal-to-clutter ratio by a factor equal to one half the product of band width and pulse length, while its signal-to-noise ratio is 6 decibels worse than a flat filter of equal band width. Filter B, the pulse-length aperture corrector, increases the signal-to-clutter ratio by a factor equal to the ratio of transmitted to corrected pulse lengths. Stated another way, the factor of increase in signal-to-clutter ratio is equal to one-fourth the product of transmitted pulse length and over-all band width. The signal-to-noise ratio is three decibels worse than the flat filter of equal band width.

Filter S, the second blip eliminator, increases the signal-to-clutter ratio by another three decibels, but it has a disastrous effect on the signal-to-noise ratio. The use of such a filter makes the problem of noise at least as severe as that of clutter. The reason for poor noise behavior is the presence of poles of infinite Q. An actual second blip eliminator will, of course, have poles of finite Q. As a matter of fact, poles and zeros of finite Q in the optimum filter represent a closer approach to the solution of the problem of clutter plus noise. It will be recalled that the optimum filter has a frequency characteristic given by the voltage spectrum of the received pulse divided by the power spectrum of the interference. If the interference consists of clutter and noise, as it always must, the optimum filter turns out to be one with poles of finite Q, the Q of each pole depending upon the ratio of clutter power to noise power. Thus, the use of a second blip eliminator with poles of finite Q is indicated. Of course, the zeros of filter A or B are also of finite Q. The use of a second blip eliminator may be undesirable for other reasons. There are indications that the second blip eliminator is very sensitive to changes in frequency, even causing an actual decrease in signal-to-clutter ratio.

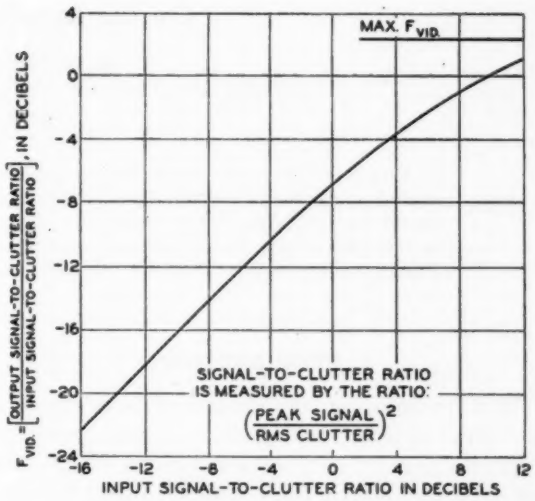


FIG. 8. Relative reduction of signal-to-clutter ratio by a linear detector.

The work described in this paper was done under the sponsorship of the Bureau of Aeronautics, Department of the Navy. The author wishes to acknowledge the generous assistance of G. J. Laurent and to express his thanks to Miss Mary Jo Roney who performed some of the computations.

On the Interference of Pulse Trains

K. S. MILLER AND R. J. SCHWARZ
New York University and Columbia University, New York, New York

(Received March 23, 1953)

A number of problems arising from the undesired overlapping or time coincidence of pulses from separate periodic pulse trains are considered. The coincidence time fraction is determined for two pulse trains with both a fixed initial phase and a randomly varying initial phase. The coincidence time fraction for these cases is then determined if only pulse coincidences equaling or exceeding a specified interval are considered. Finally, some problems involving more than two pulse trains are briefly discussed.

I. INTRODUCTION

THE problem of the undesired overlapping (in time) of pulses belonging to separate pulse trains occurs in a number of detection, control, and measurement systems utilizing pulse trains. In many cases the coincidence of pulses from two pulse trains is the desired measure of an event, but any interference of other pulse trains may make this measure unreliable. In other cases only interference beyond a certain critical amount is undesirable.

The purpose of this paper is to analyze such systems and derive certain compact formulas under rather general conditions. Several specific problems may be formulated. Figure 1 illustrates two pulse trains (abbreviated PT) of periods T_1 and T_2, respectively, pulse widths p_1 and p_2, respectively, and a time difference (here called initial phase) denoted by s. Some of the problems of interest are the following.

(a) What average fraction of time are the pulses of two PT coincident for a fixed initial phase? This is called the *coincidence problem.*

(b) What average fraction of time are the pulses of two PT coincident for a random initial phase? This is called the *random-phase coincidence problem.*

(c) What average fraction of time are the pulses of two PT coincident during a time interval equaling or exceeding a specified length if the initial phase is fixed? This is called the *tolerance coincidence problem.*

(d) What average fraction of time are the pulses of two PT coincident during a time interval equaling or exceeding a specified length if the initial phase is random? This is called the *random-phase tolerance coincidence problem.*

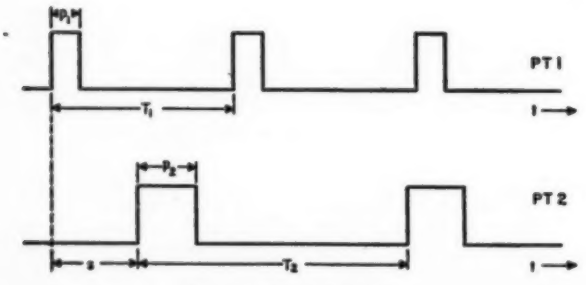


FIG. 1. Two pulse trains.

(e) How may the methods used to solve the above problems be extended to more than two pulse trains?

In the solution of these problems, use will be made of the theory of linear congruences. Some results of this theory are therefore presented in the following section as a preliminary to the consideration of the various pulse coincidence problems.

II. LINEAR CONGRUENCES

The theory of linear congruences has its roots in the pure mathematical discipline of the theory of numbers. For the remainder of this section we shall deal exclusively with *integers.*

Two integers a and b are said to be *congruent modulo* p if p divides $a-b$. This is written

$$a\equiv b \pmod{p}$$

or more simply,

$$a\equiv b \quad (p).$$

For example, $7\equiv 3$ (2) and $9\not\equiv 4$ (3) where the symbol "$\not\equiv$" is read "incongruent to" and means that 3 does *not* divide 9−4. The notation $a|b$ is used to mean "a divides b," for example, $3|12$; and the notation $a\nmid b$ is used to mean "a does not divide b," for example, $5\nmid 13$. By $d=gcd\ (a, b)$ we mean "d is the greatest common divisor of a and b," and by $M=lcm\ (a, b)$ we mean "M is the least common multiple of a and b." For example, $3=gcd\ (9, 12)$, $36=lcm\ (9, 12)$. If a and b are two integers we have the trivial identity,

$$ab=[gcd\ (a, b)][lcm\ (a, b)]. \tag{1}$$

A *linear congruence* is a congruence of the form

$$mx\equiv r \quad (n),$$

and the problem is to determine an integer (or integers) x such that $n|(mx-r)$ where m, r and n are given integers. For example, in the linear congruence

$$7x\equiv 2 \quad (3), \tag{2}$$

$x=2$ is a solution. Additional solutions are clearly given by $x=2+3k$, where k is any integer, since

$$7(2+3k)-2=12+21k=3(4+7k)$$

and $3|3(4+7k)$.

Not every linear congruence has a solution. For ex-

ample, the congruence

$$7x \equiv 1 \quad (7), \tag{3}$$

has no solutions. For if x is any integer, $7 \mid 7x$ and hence $7 \nmid (7x-1)$.

The fundamental theorem for linear congruences is the following.[1] The congruence

$$mx \equiv r \quad (n), \tag{4}$$

has a solution if $gcd\ (m, n) \mid r$ and has no solution if $gcd\ (m, n) \nmid r$.

When this theorem is applied to Eq. (2),

$$gcd\ (m, n) = gcd\ (7, 3) = 1$$

and $1 \mid 2$. Hence the congruence $7x \equiv 2$ (3) has a solution. On the other hand, for Eq. (3), $gcd\ (7,7) = 7$ and $7 \nmid 1$. Thus the congruence $7x \equiv 1$ (7) has no solution.

If one solution of a linear congruence is known, additional solutions may be located as follows. Suppose that the congruence

$$mx \equiv r \quad (n), \tag{4}$$

has solutions, and let x_0 be the smallest non-negative solution. Then $x_0 + x'$ is also a solution where

$$mx' = lcm\ (m, n). \tag{5}$$

This follows from the equation

$$m(x_0 + x') - r = (mx_0 - r) + mx' \equiv 0 \quad (n),$$

since $n \mid (mx_0 - r)$ and $n \mid mx'$. There exist no solutions between x_0 and $x_0 + x'$ as can be seen from the following argument. Suppose $x_0 + x''$ were such a solution with

$$x_0 < x_0 + x'' < x_0 + x'. \tag{6}$$

Then n must divide

$$m(x_0 + x'') - r = (mx_0 - r) + mx''.$$

Since $n \mid (mx_0 - r)$ by assumption, n must divide mx''. But $m \mid mx''$, and if both m and n are to divide mx'', the least common multiple of m and n must also divide mx''. In view of Eqs. (5) and (6) this is impossible.

III. A SPECIAL CASE

In this section a special case of two pulse trains of equal pulse width will be considered. The problems stated in Sec. I will then be solved by essentially reducing them to this case.

Figure 2 illustrates two PT of equal pulse width $p (= p_1 = p_2)$ but with different periods. It will be assumed that the periods T_1 and T_2 and the phase s are integral multiples of p, namely, $T_1 = mp$, $T_2 = np$, and $s = rp$. Coincidence can therefore occur only for a full pulse width of p seconds. These coincidences will occur, however, only when a multiple of $T_1 = mp$ exceeds a multiple of $T_2 = np$ by the initial phase $s = rp$; that is, coincidence will occur only if integers k and x exist such that

$$rp + k(np) = x(mp).$$

In congruence notation, coincidence will occur only if the congruence

$$mx \equiv r \quad (n), \tag{7}$$

has a solution.

If Eq. (7) is assumed to have a solution, then, as shown in the previous section, coincidences will occur at intervals of $mx'p$ seconds where

$$mx' = lcm\ (m, n).$$

Over a long period of time, pulses of the two pulse trains will coincide for a time fraction f given by

$$f = \frac{p}{mx'p} = \frac{1}{M},$$

where $M = lcm\ (m, n)$. If the congruence of Eq. (7) does not have a solution, then $f = 0$.

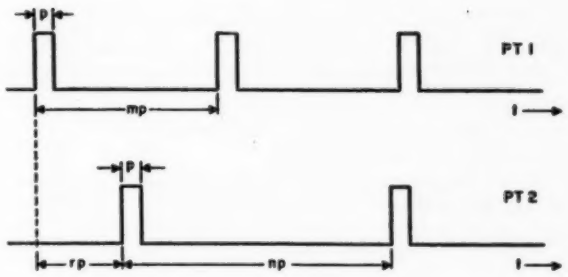


FIG. 2. Two pulse trains of equal pulse width.

IV. THE COINCIDENCE PROBLEM

Figure 1 illustrates the general case to be considered in this section. It will be assumed that p is a quantity so small that p_1, p_2, T_1, T_2, and s are all integral multiples of p:

$$p_1 = \mu p, \quad T_1 = mp,$$
$$p_2 = \nu p, \quad T_2 = np,$$
$$s = rp.$$

From a practical point of view, this is always possible. If two or more of the quantities p_1, p_2, T_1, T_2, s are incommensurable, a limiting argument may be used.

Pulse train PT 1 may be resolved into the μ-pulse trains PT 1_1, PT 1_2, $\cdots$, PT 1_μ shown in Fig. 3. Each PT 1_j has a pulse width of p seconds, a period mp, and a phase (relative to PT 1_1) of $(j-1)p$ seconds. Similarly PT 2 may be resolved into ν pulse trains PT 2_1, PT 2_2, $\cdots$, PT 2_ν. The occurrence of pulse coincidences of the trains PT 1_j and PT 2_k for $j = 1, 2, \cdots, \mu$, $k = 1, 2, \cdots, \nu$ may now be analyzed in terms of the equal-width PT of Sec. III. The sum of the coincidence time of each of the $\mu\nu$ pairs made up of one component of PT 1 and one component of PT 2 is the total coincidence time of the two PT of Fig. 1.

[1] L. E. Dickson, *Introduction to the Theory of Numbers* (The University of Chicago Press, Chicago, 1929).

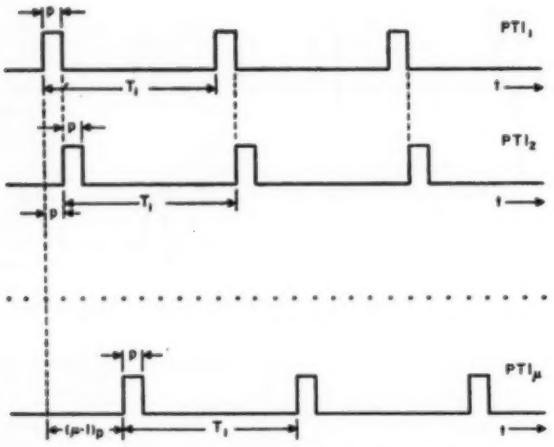


FIG. 3. Decomposition of a pulse train.

The component PT 1_1 will coincide with the component PT 2_1 the fraction $1/M$ [where $M=lcm\ (m, n)$] if the congruence

$$mx\equiv r \quad (n),$$

has a solution, and will not coincide if this congruence does not have a solution. Similarly, PT 1_2 will coincide with PT 2_1 the fraction $1/M$ if the congruence

$$mx\equiv r-1 \quad (n),$$

has a solution, and will not coincide otherwise. Continuing in this manner, it is seen that PT 1 coincides with PT 2_1 the fraction

$$f_1=\phi_1/M,$$

where ϕ_1 is the number of the congruences

$$mx\equiv r-i \quad (n), \quad i=0, 1, \cdots, \mu-1,$$

which have solutions.

Similarly, PT 1 will coincide with PT 2_2 the fraction

$$f_2=\phi_2/M,$$

where ϕ_2 is the number of the congruences

$$mx\equiv r+1-i \quad (n), \quad i=0, 1, \cdots, \mu-1,$$

which have solutions.

It is thus seen that PT 1 will coincide with PT 2_α the fraction

$$f_\alpha=\phi_\alpha/M,$$

where ϕ_α is the number of the congruences

$$mx\equiv r+\alpha-1-i \quad (n), \quad i=0, 1, \cdots, \mu-1, \qquad (8)$$

which have solutions. Hence the original pulse trains PT 1 and PT 2 will coincide the fraction

$$f^r=\sum_{\alpha=1}^{\nu} f_\alpha=\frac{1}{M}\sum_{\alpha=1}^{\nu}\phi_\alpha, \qquad (9)$$

where the superscript r refers to the initial phase rp. Our first problem is therefore solved.

An equivalent expression for the coincidence fraction may be obtained by considering the coincidence of PT 2 with PT 1_1, PT 1_2, $\cdots$, PT 1_μ. One then arrives at the formula

$$f^r=\frac{1}{M}\sum_{\beta=1}^{\mu}\phi^\beta,$$

where ϕ^β, $\beta=1, 2, \cdots, \mu$ is the number of the congruences

$$mx\equiv r+\beta-1-j \quad (n), \quad j=0, 1, \cdots, \nu-1,$$

which have solutions.

V. THE RANDOM-PHASE COINCIDENCE PROBLEM

If PT 2 represents undesired interference with PT 1, the initial phase is generally not known and may reasonably be assumed to be random with a rectangular distribution. If separate interferences of PT 2 with PT 1 occur over a long interval of time, the pulse-coincidence fraction is then obtained by considering the average value of f^r over all possible initial phases $s=rp$. It is sufficient to consider

$$r=0, 1, \cdots, h-1,$$

where $h=gcd\ (m,n)$, since $r=h$ again corresponds to zero initial phase between the two PT.

It is thus necessary to find the number of the congruences

$$mx\equiv z_r \quad (n), \qquad (10)$$

which have solutions, where z_r is any element of the matrix Z_r,

$$Z_r=\begin{Vmatrix} r & r+1 & \cdots & r+\nu-1 \\ r-1 & r & \cdots & r+\nu-2 \\ r-2 & r-1 & \cdots & r+\nu-3 \\ \cdot & \cdot & \cdot & \cdot \\ r-\mu+1 & r-\mu+2 & \cdots & r+\nu-\mu \end{Vmatrix},$$

and where r takes on the values $0, 1, \cdots, h-1$.

Every element z_r of the matrix Z_r may be written

$$z_r=Ah+(b+r) \quad r=0, 1, \cdots, h-1,$$

where A is some integer and b is a non-negative integer satisfying the inequalities

$$0\leqq b<h.$$

For any one element whose value of b is fixed, one and only one of the possible values of r is such that $b+r\equiv 0$ (h). Hence as r varies from 0 to $h-1$, each element z_r becomes divisible by h exactly once. In view of the congruence theorem stated in Sec. II, it is seen that, among the $h(\mu\nu)$ congruences of Eq. (10) represented by the elements of the h matrices Z_r, $r=0, 1, \cdots, h-1$, exactly $\mu\nu$ have solutions. Thus

$$f=\frac{1}{h}\sum_{r=0}^{h-1} f^r=\frac{1}{hM}(\mu\nu)=\frac{\mu\nu}{mn},$$

since $hM=mn$ by Eq. (1).

The fraction f may be written

$$f=\frac{(\mu p)(\nu p)}{(mp)(np)}=\frac{p_1p_2}{T_1T_2}, \tag{11}$$

and clearly the limit of f as p approaches zero is independent of p.

Summarizing this, if PT 1 and PT 2 are two PT of pulse widths p_1 and p_2, respectively, and periods T_1 and T_2, respectively; then, on the average, PT 1 and PT 2 will coincide, in a long period of time, the time fraction

$$f=\frac{p_1p_2}{T_1T_2}.$$

This result could also be obtained by probability considerations.

VI. THE TOLERANCE COINCIDENCE PROBLEM

In some applications, PT are applied to systems biased in such a manner that pulses from the second PT must overlap with pulses from the first PT for a specified minimum time interval before triggering occurs. A third undesired interfering PT can thus cause spurious triggering only if it coincides with the first PT for at least this specified time interval. In order to determine the interference probability, it is thus necessary to determine the fraction of time for which coincidences equaling or exceeding a specified duration occur.

Returning to the two PT of Fig. 1, it may be assumed without loss of generality that $\mu\leqq\nu$; then only coincidence of widths p, $2p$, $\cdots$, μp can occur. The pulse width μp of PT 1 will be interior to the pulse width νp of PT 2 (resulting in full coincidence) if

$$rp+k(np)=x(mp)-ip, \quad i=0, 1, \cdots, \nu-\mu,$$

for some integers k and x (depending on i). In congruence notation, these equations become

$$mx\equiv r+i \quad (n), \quad i=0, 1, \cdots, \nu-\mu.$$

If ψ_μ is the number of these congruences which have solutions, the fraction of time $g_\mu{}^r$ for which coincidence of width μp occurs, when the initial phase is rp, is

$$g_\mu{}^r=\frac{\mu\psi_\mu}{M}. \tag{12}$$

The multiplicative factor μ appears since each coincidence is of width μp.

The number of "partial" coincidences of width $(\mu-\zeta)p$, $\zeta=1, 2, \cdots, \mu-1$ is given by the number $\psi_{\mu-\zeta}$ of the congruences

$$mx\equiv r-\zeta \quad (n),$$
$$mx\equiv r+\nu-\mu+\zeta \quad (n),$$

which have solutions. Coincidences of width $(\mu-\zeta)p$ when the initial phase is rp thus occur for a fraction of time

$$g_{\mu-\zeta}{}^r=\frac{(\mu-\zeta)\psi_{\mu-\zeta}}{M}. \tag{13}$$

Pulse trains PT 1 and PT 2 with initial phase rp are thus coincident for an interval equaling or exceeding αp, $0<\alpha\leqq\mu\leqq\nu$ for a fraction of time

$$g^{r,\alpha}=\sum_{k=0}^{\mu-\alpha} g_{\mu-k}{}^r=\frac{1}{M}\sum_{k=0}^{\mu-\alpha}(\mu-k)\psi_{\mu-k}. \tag{14}$$

Equation (14) solves problem (c).

VII. THE RANDOM-PHASE TOLERANCE COINCIDENCE PROBLEM

If separate interferences of PT 2 with PT 1 occur over a long period of time during which the initial phase varies at random, the average coincidence fraction may be obtained, as in Sec. V, by considering the average value of $g^{r,\alpha}$ over all possible phases $s=rp$.

Averaging the $g_\mu{}^r$ of Eq. (12) over all phases r, it is found as in Sec. V that

$$g_\mu=\mu\frac{1}{h}\frac{1}{M}(\nu-\mu+1)=\frac{\mu(\nu-\mu+1)}{mn}, \tag{15}$$

while averaging the fraction $g_{\mu-\zeta}{}^r$ of Eq. (13) yields

$$g_{\mu-\zeta}=\frac{2(\mu-\zeta)}{mn}, \quad \zeta=1, 2, \cdots, \mu-1. \tag{16}$$

(It is noted that the average fraction $f=\mu\nu/mn$ must equal the sum of the g_k's. This is readily verified by a direct calculation.)

The average fraction F_α of time for which the coincidence width equals or exceeds αp, $0<\alpha\leqq\mu\leqq\nu$ is determined by adding the fractions representing coincidences for αp and any longer duration. Clearly

$$F_\alpha=\sum_{k=0}^{\mu-\alpha} g_{\mu-k}=\frac{\mu(\nu-\mu+1)}{mn}+\frac{2(\mu-1)}{mn}+\frac{2(\mu-2)}{mn}+\cdots+\frac{2\alpha}{mn}$$
$$=\frac{\mu\nu-\alpha^2+\alpha}{mn}. \tag{17}$$

If the original pulse widths and periods are re-introduced, $p_1=\mu p$, $p_2=\nu p$, $T_1=mp$, $T_2=np$, and if further

$$u=\alpha p,$$

then Eq. (17) may be written as

$$F_\alpha=\frac{p_1p_2-u^2+pu}{T_1T_2}.$$

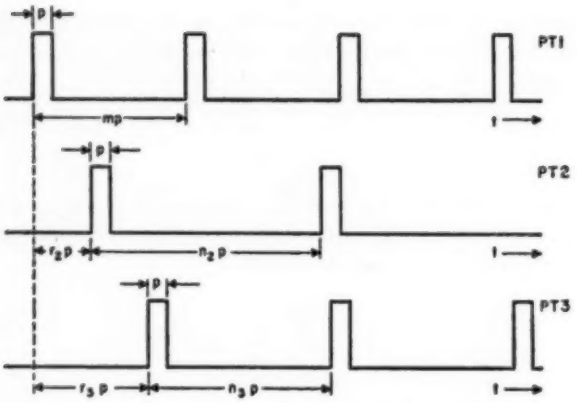


FIG. 4. Three pulse trains of equal pulse width.

The limit of this expression as the elementary pulse width p approaches zero is

$$F_u=\frac{p_1p_2-u^2}{T_1T_2}. \tag{18}$$

Summarizing this, if PT 1 and PT 2 are two PT of pulse widths p_1 and p_2, respectively, and periods T_1 and T_2, respectively; then, on the average, PT 1 and PT 2 will coincide, in a long period of time, for a time greater than or equal to u seconds, the time fraction

$$F_u=\frac{p_1p_2-u^2}{T_1T_2},$$

where $0\leqq u\leqq p_1\leqq p_2$.

If $u=0$, that is, if coincidences of any width are allowed, then Eq. (18) reduces to Eq. (11).

VIII. COINCIDENCES INVOLVING MORE THAN TWO PULSE TRAINS

In some systems, interference may be caused simultaneously by several pulse trains PT 2, PT 3, $\cdots$ and the likelihood of the occurrence of such interference may be determined from the fraction of time for which any one or any combination of the interfering pulses coincides with pulses of PT 1. If pulse coincidence of one train, say PT 2 with PT 1 is the desired measure of an event, the coincidences of PT 3, PT 4, $\cdots$ with PT 1 are not harmful during any intervals in which PT 2 also coincides with PT 1.

The total interference of two pulse trains PT 2 and PT 3 with a pulse train PT 1 may be determined as

$$f=f^{12}+f^{13}-f^{123},$$

where

$f^{\alpha\beta\cdots\lambda}$ = fraction of time for which PT α, PT β, $\cdots$, PT λ simultaneously coincide.

One can calculate f^{12} and f^{13} by the previously developed techniques. The fraction f^{123} can be determined as follows:

If the PT all have pulse widths p (see Fig. 4), then all three PT will coincide if

$$A(mp)=r_2p+B(n_2p)=r_3p+C(n_3p),$$

where A, B, and C are some integers. In congruence notation, if the *simultaneous* congruences

$$mx\equiv r_2 \quad (n_2),$$
$$mx\equiv r_3 \quad (n_3),$$

have a solution for the *same* x, then PT 1, PT 2, and PT 3 will coincide.

These results can be generalized to PT of arbitrary pulse width. For the case of n PT, the total interference of PT 2, PT 3, $\cdots$, PT n with PT 1, is

$$f=\sum_{\alpha=1}^{n}\sum_{j=1}^{\binom{n}{\alpha}}(-1)^{\alpha+1}f^{1,\sigma_j{}^{\alpha}},$$

where $\sigma_j{}^{\alpha}$ is one of the $\binom{n}{\alpha}$ combinations of the $n-1$ numbers 2, 3, $\cdots$, n taken α at a time. For example, for $n=5$,

$$f=(f^{12}+f^{13}+f^{14}+f^{15})-(f^{123}+f^{124}+f^{125}+f^{134}+f^{135}+f^{145})+(f^{1234}+f^{1235}+f^{1345})-f^{12345}$$

These results can be deduced either by considering the algebra of classes or by probability theory.

Beta-Ray Spectrometer Line Shapes with Tilted Sources*

H. W. KRUSE, G. P. MELLOR, AND C. M. FOWLER
Kansas State College, Manhattan, Kansas

(Received March 23, 1953)

Experimental line shapes arising from tilted beta-ray spectrometer sources are compared with analytically derived line shapes. Although the increased scattering in the source at higher tilting angles considerably alters the experimental line shape, the sharp leading edges predicted for the larger angles are largely retained. This is true for electron energies at least as low as 400 kev, even for quite dense sources (5 mg/cm²). An application of the increased resolution resulting from this feature is given.

INTRODUCTION

IN a recent paper,[1] an analysis was presented which described theoretical beta-ray spectrometer line shapes for tilted sources. The analysis applied to the line structure obtained with a fixed field, 180° type focusing spectrometer. The source electrons were supposed monoenergetic, and emitted both isotropically and uniformly over the source plane. Figure 1 gives a view of the focusing geometry perpendicular to the magnetic field. Under normal conditions, the source plane (projected as *cd*) is parallel to the focal plane. The source is said to be tilted an angle α when rotated to the position shown in the figure.

Although the idea of source tilting is an old one,[2] it has had little application in recent years. On the other hand, the theory predicts very definite advantages for the tilted source. Line peak heights increase, line widths decrease, and the leading edge sharpens as the angle of tilting moves from zero to 90°. A major practical defect, however, is the increased source scattering as α is increased. It is the purpose of this paper to compare some experimental line shapes with the analytically predicted line shapes. Although considerable departures from the analytic line shapes are observable, it will be seen that the sharp leading edges predicted, are largely retained. An application utilizing this feature is discussed.

THEORETICAL LINE SHAPES

The essential features necessary to compute the theoretical line shapes when the electron momenta are all perpendicular to the magnetic field (called the Z direction) are shown in Fig. 1. These electrons, all having the same energy, describe arcs with the same radius R. It is readily seen that the centers of all trajectories contributing to the line structure lie inside the region A. This region is constructed by drawing four circles, radius R, centered at the slit and source limits. Those particle trajectories making up the line density at P must then all have centers which lie on the arc xy within the region A. The two angles θ_1 and θ_2 determine the intensity at the point P. In the usual situation $\alpha=0$, the intensity,[3] I_p is directly proportional to the difference of the angles, $\theta_1-\theta_2$. When $\alpha\neq 0$, the analytical expression, given in reference 1, is a rather formidable function of the two angles.

The finite heights of both source and detector (in the Z direction) allow the detector to accept electrons whose momenta are not completely perpendicular to the field. These electrons are focused in circles of smaller radii, depending upon the angle of the electrons away from the field perpendicular, and give rise to a low energy broadening of the ideal line described above. Corrections for this effect can be simply calculated by numerical integrations. In the present case, the corrections were made by plotting separate line shapes for small increments of angles away from the field perpendicular. Each line shape was weighted by the fraction of the emission which could be collected by the detector. The individual incremental lines were then added together to give the resultant line shape, corrected for Z defocusing. A final graphical integration over the detector window widths gave the theoretical line shapes.

Figure 2 shows some of these lines computed with the parameters used in the experimental investigations described below. The only parameter different for the three lines is the rotation angle α.

Before correcting the lines for Z defocusing, the 85° line showed a peak height of roughly two and a half times that for the 0° line, while its half-width was a

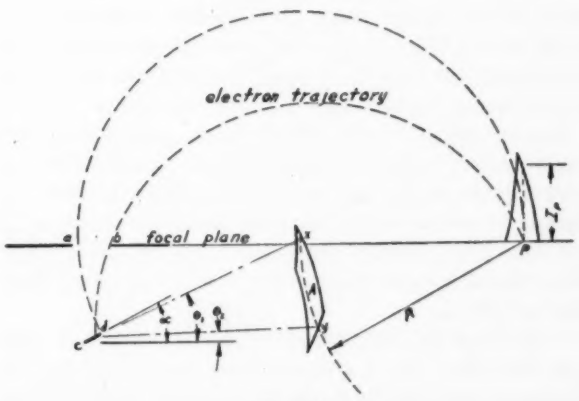


FIG. 1.

* Sponsored by the U. S. Atomic Energy Commission.

[1] C. M. Fowler and P. Domotor, J. Appl. Phys. **23**, 415 (1952).

[2] W. A. Wooster, Proc. Roy. Soc. (London) **A114**, 729 (1927).

[3] Fowler, Shreffler, and Cork, Rev. Sci. Instr. **20**, 966 (1949).

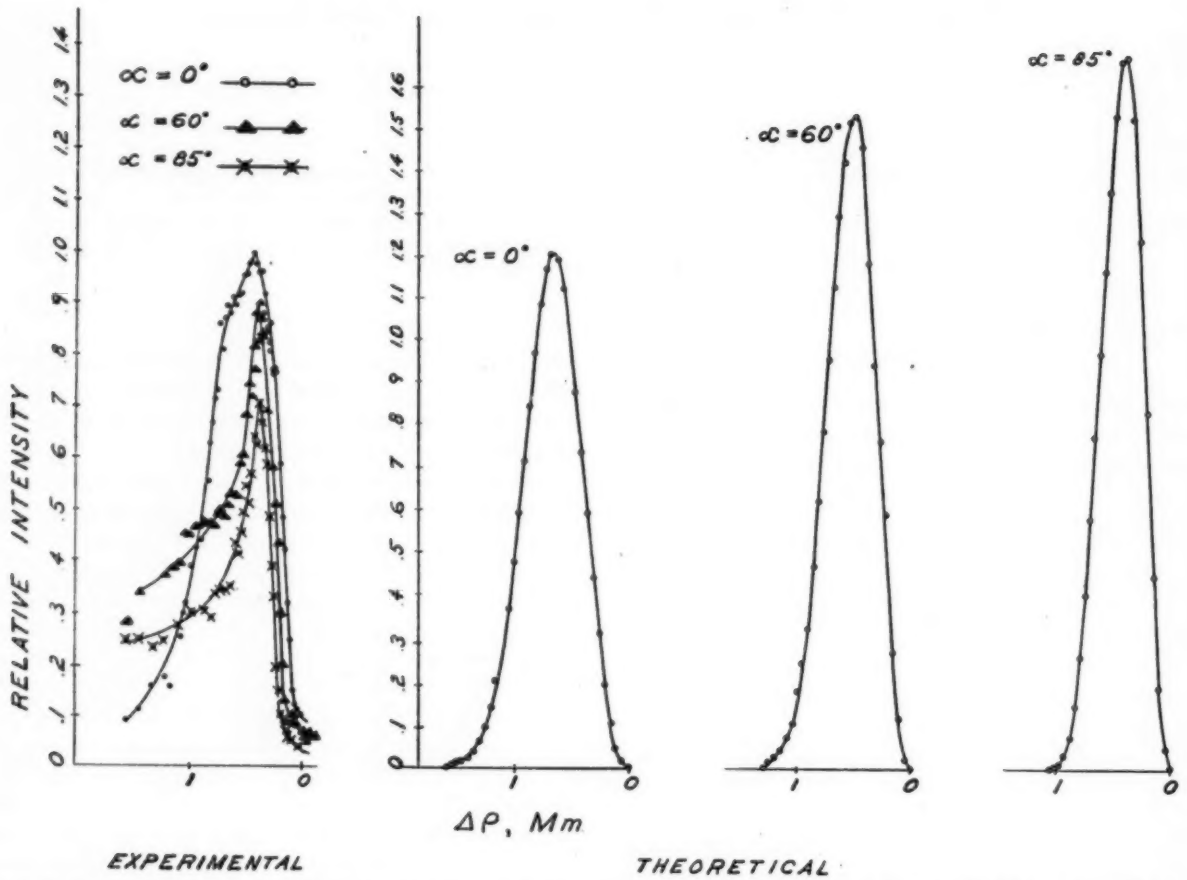


FIG. 2. Experimental and theoretical line shapes for 624-kev conversion group. Intensities are plotted against change in radius of curvature. Radii are about 8.40 cm at line peaks. Source 1 mm wide ×10 mm high, 100μ grams/cm^2, 2.10 cm from focal plane. Slits 4 mm wide. Geiger tube window 0.040 cm wide ×2.50 cm high.

little less than half of that of the 0° line. Correction for Z defocusing accounted for most of the changes of the ideal shapes to those seen in the figure.

EXPERIMENTAL LINE SHAPES

The strong K conversion line (624 kev) arising in the decay of the fission product Cs^{137} was selected for quantitative study, since high specific activities were available. Further, the electrons are of reasonably high energy, and the half-life is long. The first two factors are advantageous in minimizing source scattering.

Measurements were taken in a permanent field spectrometer, using a Geiger tube as the detector. Since significant changes of line structure occurred over changes of radius of less than one part in a thousand, it was felt that current induced fields would present difficulties of stabilization, over the long operating times often required.

A threaded spindle was attached to the Geiger tube such that the tube could traverse the line shape, by rotation of the spindle through a control knob, outside the camera. A veeder-Root counter attached to the spindle recorded the position of the tube.

It is clear from Fig. 1, that the location of the source center with respect to the slit center both along and perpendicular to the focal plane influences the line shape, since changes of these parameters alter the area A. In the present study these parameters were held fixed by mounting the source strip on an aluminum frame, whose only degree of freedom was a rotation to allow for studies at different angles of α. The axis of rotation, lying in the plane of the source, was adjusted to pass through the source midpoint of *cd*. The axis was also located symmetrically with respect to the slit edges; that is, projected vertically on the focal plane, the axis would lie midway between the slits.

To keep Z defocusing as small as possible, both source and detector heights were made smaller than those conventionally employed. Geiger tube window widths were also considerably narrower than those normally used. These, and other significant parameters are given in the caption of Fig. 2, which compares the experimental and theoretical line shapes.

Within the limits of accuracy, the ratios of areas under the three experimental curves for which $\alpha=0°$,

60°, 85°, agreed with those for the corresponding calculated curves. This is evidence that the electrons were mainly scattered from the beam but not removed from it. Consequently, the intensity scale of the theoretical curves was adjusted to give the same area under the 0° curve, as that observed experimentally for this angle. The peak height of the experimental curve for 0° was arbitrarily set at unity.

The increased scattering with increasing angle, is quite apparent in the large low energy tails seen at larger values of α. In spite of this, the half-widths of the large α lines are substantially better than for the line, $\alpha=0°$. The higher peak heights expected, fail to materialize, apparently being lost into the low energy tail.

A prominent feature of the larger α lines is the retention of the sharp leading edge. This is somewhat puzzling in view of the evident source scattering.

To see how many of these features were retained under more adverse source conditions, a Ta^{182} source of approximately 5 mg/cm² was studied in the neighborhood of the strong 1051 kev internal conversion line. Figure 3 compares the line for $\alpha=0°$ with that at $\alpha=80°$. Although the half-width and peak height fall far short of theoretical predictions, the sharp leading edge at the larger angle is retained.

A faint line at about 1060 kev, previously observed photographically, was completely obscured in the zero-angle study. The resolution of this line, in the 80° study, shows that application of the tilted source technique is sometimes advantageous.

Thick source studies, similar to those described above for Ta, were also carried out with conversion groups of energies as low as 400 kev. The slope of the leading edges for high angles is lessened somewhat but is still greater than that for no rotation.

The sharp leading edges arising from tilted sources also offer an advantage in photographic spectroscopy. The contrast with background offered by such edges facilitates the observation of very faint lines. When this technique is used, a number of very weak radiations have been observed in this laboratory.

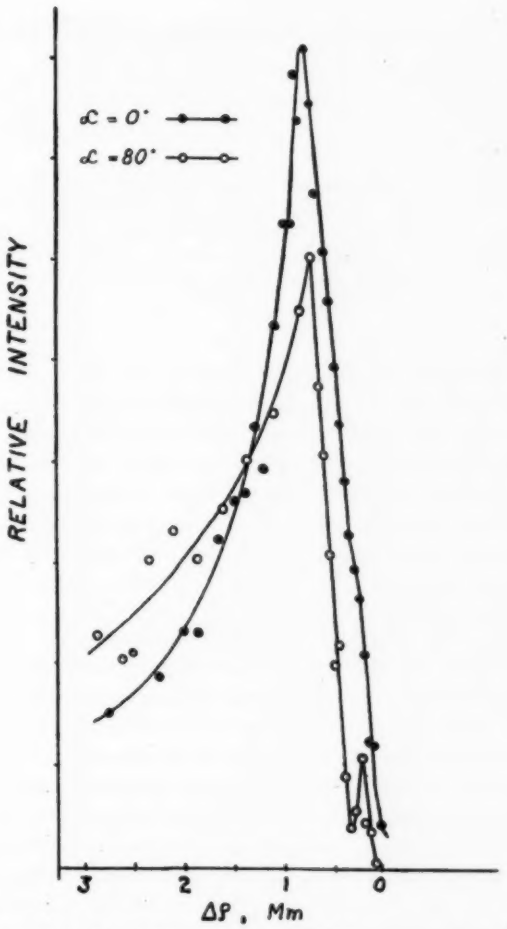


FIG. 3. Experimental line shapes for 1051-kev conversion groups. Intensities are plotted against change in radius of curvature. Radii are about 8.90 cm at line peaks. Source 1.5 mm wide × 10 mm high, 5 mg/cm², 2.50 cm from slit plane. Slits 6 mm wide. Geiger tube window 0.050 cm wide × 1.60 cm high.

Measurement of Particle Size of Fine Ferromagnetic Powders*

A. D. FRANKLIN, R. CAMPBELL, AND J. WEINMAN†
The Franklin Institute Laboratories for Research and Development, Philadelphia, Pennsylvania
(Received March 26, 1953)

The particle sizes of four Fe and two Fe-Co powders with diameters ranging from 200 to 600 angstroms have been determined using electron microscope, x-ray line broadening, and nitrogen adsorption techniques. Reasonable agreement was found between the electron microscope and the x-ray methods, while the sizes determined through nitrogen adsorption were in every case larger. A method of dispersing single-domain ferromagnetic powders for electron microscopy is described.

UNDER certain circumstances, the measurement of the sizes of submicron particles can be very difficult. Except in those favorable cases when sols[1] can be formed, methods such as sedimentation, elutriation, etc. are useless because they measure clump rather than particle size. If the material is ferromagnetic, this problem is even more severe, since the particles tend to contain but one ferromagnetic domain, and to behave as tiny permanent magnets.

The best particle size methods with such powders appear to be the use of the electron microscope, the determination of surface areas through nitrogen adsorption, and the measurement of the broadening of x-ray diffraction lines. Use of the electron microscope requires that the powders be reasonably well dispersed, while the x-ray method should be almost independent of, and the surface area method somewhat dependent upon, clumping. The electron microscope is capable of yielding the complete particle size distribution in a direct and simple fashion. The x-ray method, except under very favorable experimental conditions,[2,3] and the surface area method give specific averages, which may not be applicable to the particular problem of interest. In this report, the electron microscope is considered to be the central technique. Comparison is made with the results obtained by the other two methods in order to check upon the microscope.

A number of studies using one or two of these methods for nonferromagnetic materials[1,4–7] and for ferromagnetic materials[8,9] have been reported. In most of the studies involving the use of x-ray diffraction line broadening, however, approximate methods for correction for instrumental broadening were used. Furthermore, no study of all three methods simultaneously has been reported to the authors' knowledge. Finally, there appears to be no comprehensive report in the literature of the use of the electron microscope with single-domain ferromagnetic particles.

For these reasons, a study of the particle size determination of fine iron and iron-cobalt powders with diameters in the range from 100 angstroms to 2 microns has been made, using all three of the methods mentioned above.

TABLE I. Properties of powders used.

Designation[a]	Size	Composition	Percent free metal
GAF HP	5 microns	Fe	99
GAF P-818	2 microns	Fe	98
Fe-8	600 angstroms	Fe	87
Fe-7	400 angstroms	Fe	69
Hyflux	400 angstroms	Fe	90
NOL-1	250 angstroms	Fe	79
NOL-2	200 angstroms	Fe-Co	40
Ugine	200 angstroms	Fe-Co	58

[a] HP, P-818 and Ugine were supplied by the General Aniline and Film Corporation, Hyflux by the Indiana Steel Products Company, and NOL-1 and NOL-2 by the Magnetics Laboratory of the U. S. Naval Ordnance Laboratory.

MATERIALS

The approximate sizes and compositions of the powders used are listed in Table I.

HP and P-818 are carbonyl iron; Hyflux was produced by the Raney process; the rest were made by reducing the appropriate metal formates. All of the finer powders contain a considerable amount of oxide.

EXPERIMENTAL

1. Electron Microscope

Micrographs of the powders embedded in Formvar or collodion films were obtained with an RCA EMU-A microscope. The specimens for the larger powders were prepared by simple grinding of suspensions of the powders, spreading the suspension between glass slides to form the films, and transferring them to copper

* This work was supported by the U. S. Office of Naval Research. Papers based upon this material were presented at the Tenth Annual Meeting of the Electron Microscope Society of America, Cleveland, Ohio, November 6, 1952; and at the Durham, North Carolina, Meeting of the American Physical Society in March, 1953.

† Present address: University of Illinois, Urbana, Illinois.

[1] Turkevich, Stevenson, and Hillier, Disc. Faraday Soc. **11**, 55–75 (1951).

[2] B. E. Warren and L. Averbach, J. Appl. Phys. **21**, 595 (1950).

[3] F. Bertaut, Acta Cryst. **3**, 14 (1950).

[4] U. Hofmann and D. Wilm, Z. physik. Chem. **B18**, 401 (1932), Z. Elektrochem. **42**, 504 (1936).

[5] J. Biscoe and B. E. Warren, J. Appl. Phys. **13**, 364 (1942).

[6] A. C. Zettlemoyer and W. C. Walker, J. Phys. and Colloid Chem. **51**, 763 (1947).

[7] L. S. Birks and H. Friedman, J. Appl. Phys. **17**, 687 (1946).

[8] Z. W. Wilchinsky, Acta Cryst. **4**, 1 (1951).

[9] B. Kopelman, Elec. Eng. **71**, 447 (1952); Y. Schubert and B. Kopelman, Powder Met. Bull. **6**, 105 (1952).

screens with Scotch tape. For the finest powders, however, the grinding failed to give good separation of the particles, probably because of the strong magnetic forces involved.

In order to obtain usable micrographs with the very fine powders, a magnetic dispersion technique was used. The powder was dispersed in a very viscous collodion-amyl acetate mixture by exposing it to an ac magnetic field of the order of 1800 oersteds rms. The suspension was contained in a gelatin capsule, which was placed in a small gap cut in the core of an ordinary radio filter choke. Sufficient current was passed through the choke to produce the ac field. A glass ball was placed in the capsule, and the whole assembly rocked slowly so that the glass ball rolled from end to end in the capsule. The films were formed from these suspensions in the same fashion as above, after exposure to the ac field for periods from several hours to a week. Apparently, the efforts of the particles to follow the ac field in the viscous medium forced them apart. This, combined with the mechanical mixing owing to the glass ball, produced reasonably good dispersions. In Fig. 1, the dispersions produced both by the grinding and the magnetic techniques for the Ugine powder are compared, and in Fig. 2 some representative micrographs of other powders are shown.

It was not found possible to completely eliminate the clumps with any of the materials. However, enough small groups and single particles could be found to make reasonable size counts. It was then necessary to assume that the distributions of particles in uncountable clumps was the same as among those counted, so that the counts could be taken as representative of the whole. The agreement found among the three methods indicates that this is a reasonable assumption.

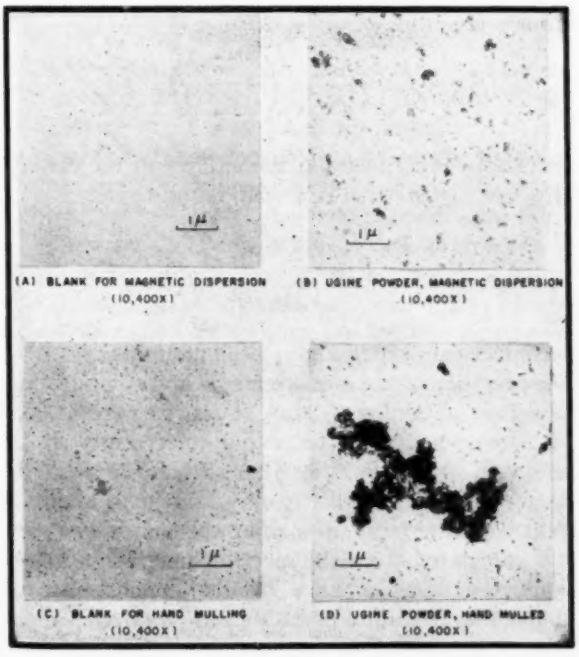


FIG. 1. Electron micrographs of Ugine powders.

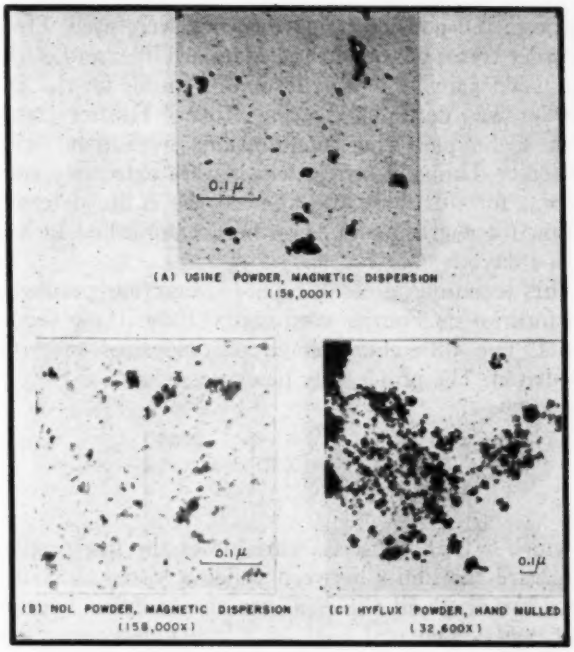


FIG. 2. Electron micrographs of Ugine, NOL, and Hyflux powders.

The microscope was calibrated with replicas of a 15 000 line per inch diffraction grating. The estimated probable error is of the order of 10 percent.

2. X-Ray Diffraction Line Broadening

A General Electric Geiger Counter Spectrometer Model XRD-3 was used to obtain the principal (110) diffraction line for all samples. The lines were obtained by determining counts per second with the Geiger counter set manually. Using an iron target with a wavelength of 1.934 angstroms, this line occurs at $2\theta = 57°$. The $\alpha_1 - \alpha_2$ doublet was not resolved.

Broadening of x-ray diffraction lines can be caused by small crystals, lattice strains, and stacking faults. In addition, a certain breadth of the line is due strictly to the instrument. Once the instrumental breadth has been accounted for, the problem arises of differentiating among the three causes of line breadth in the pure diffraction profile. Warren and Averbach[2] have developed a method of distinguishing among them, which unfortunately requires obtaining several lines in the x-ray pattern. With the present apparatus, only the most intense line was sufficiently strong to be used. It was therefore necessary to treat all line broadening as due to particle size alone. Considering the history of these particles, this is not too unreasonable an assumption.

In order to account for the broadening due to the instrument, the same line was obtained from iron samples containing particles larger than one micron for which no particle size broadening occurs. For this

purpose, HP powder and Armco iron were used. From recorder traces of one of these standard lines and of the fine powder line, the pure diffraction profile for the fine powder was determined using Stokes[10] Fourier transform technique. The computations were done with Patterson-Tunnell[11] strips, which were extremely convenient for this task, and allowed the entire determination for a given sample to be accomplished in less than a day.

This technique gives the pure diffraction profile in the form of its Fourier coefficients. From these coefficients, two different average particle sizes may be calculated. The profile may be expressed as

$$f(x)=\sum_{-\infty}^{\infty} F(n)\exp\left[-i\frac{2\pi nx}{a}\right]$$

in which x is the abscissa variable of the line profile, $\pm a/2$ are the limits between which x varies, n is the Fourier index, and the $F(n)$ are the complex Fourier coefficients. The real part of $F(n)$ is designated by $F_r(n)$ and the imaginary part by $F_i(n)$.

The integral breadth of the line is then given by

$$\beta=\frac{\text{Area}}{\text{Maximum amplitude}}=\frac{aF_r(0)}{F_r(0)+2\sum_1^{\infty}F_r(n)}, \quad (1)$$

where use has been made of the symmetry properties, $F_r(n)=F_r(-n)$ and $F_i(n)=-F_i(-n)$. Equation (1) assumes that the line peak occurs at $x=0$, for which a necessary condition is

$$\sum_1^{\infty}F_i(n)=0.$$

This equation represents a reasonable criterion of the applicability of Eq. (1) to the data at hand. From this integral breadth, an average particle size may be calculated using the Scherrer equation[12]

$$d_\beta=\frac{K\lambda}{\beta\cos\theta_0}. \quad (2)$$

Here K is the Scherrer constant, assumed to have a value of 0.94[11]; λ is the x-ray wavelength; β the integral breadth; and θ_0 the Bragg angle at the peak.

This average size may also be calculated from the particle size distribution given by the electron microscope, for comparison with the x-ray value. Bertaut[3] has shown that this average is given by

$$d_\beta=\frac{\int\int M^2 d\xi_1 d\xi_2}{\int\int M d\xi_1 d\xi_2}$$

in which M is the particle dimension perpendicular to the diffracting plane, and ξ_1, ξ_2 are the spatial coordinates in this plane. For spherical particles, this reduces to

$$d_\beta=\frac{3\sum n_i d_i^4}{4\sum n_i d_i^3}, \quad (3)$$

where n_i is the number of particles in the size range with average diameter d_i. Jones[13] suggested the same average without the factor of $\frac{3}{4}$.

Warren and Averbach[8] have shown that if a plot of $F_r(n)$ against n be made, then the slope of the curve at $n=0$ is the inverse of the average number of "unit cells" in a column perpendicular to the diffracting plane:

$$\bar{L}=\left(\frac{\partial F_r(n)}{\partial n}\right)_{n=0}^{-1},$$

while the "unit cell" dimension a_3 is

$$a_3=\frac{\lambda}{a\cos\theta_0},$$

where the other quantities have the same significance as above. Then the average length of the perpendicular column is given by

$$d_s=\bar{L}a_3=\frac{\lambda}{a\cos\theta_0}\left(\frac{\partial F_r(n)}{\partial n}\right)_{n=0}^{-1}. \quad (4)$$

The corresponding average diameter from the electron microscope distribution is readily found to be

$$d_s=\frac{2\sum n_i d_i^3}{3\sum n_i d_i^2}, \quad (5)$$

where spherical particles have been assumed.

The precision of the measurements and computations involved in obtaining d_β was about 10 percent, and for d_s, somewhat worse. A major uncertainty in the accuracy in d_β lies in the value of the Scherrer constant K, in Eq. (2). This is only known to about 20 percent.[14] Furthermore, in making the comparison with values calculated from the electron microscope distributions, the particles have been assumed spherical, although they are quite probably somewhat ellipsoidal.

[10] A. R. Stokes, Proc. Roy. Soc. (London) **61**, 382 (1948).
[11] A. L. Patterson and G. Tunnell, J. Min. Soc. Am. **27**, 655 (1942).
[12] R. W. James, *The Optical Principles of the Diffraction of X-Rays* (G. Bell and Sons, London, 1948), p. 536.
[13] F. W. Jones, Proc. Roy. Soc. (London) **166A**, 16 (1938).
[14] L. Alexander and H. P. Klug, J. Appl. Phys. **21**, 137 (1950).

Taken together, these uncertainties do not admit of an accuracy better than about 30 percent in the x-ray data.

3. Surface Area Measurements

The surface areas were obtained from nitrogen adsorption isotherms at 77°K using the BET[15] technique. The apparatus was of the usual volumetric kind.[16,17] The nitrogen used was Mathieson prepurified, stated to be 99.9 percent pure. It was passed through a drying tube containing Ascarite and Drierite, and then over copper filings maintained at 500°C. Mathieson helium, supplied at 99.8 percent pure, was used to measure the dead space volume and the true density of the material.

The samples were outgassed at about 200° in a pressure of less than 1 micron of mercury for two to three hours in agreement with Emmett's[16] recommendations. These conditions were chosen after making a series of runs with different outgassing temperatures. In Fig. 3, the surface area and the parameter c from the BET equation for these isotherms are plotted against the outgassing temperatures. The decrease in the surface area with increasing outgassing temperature was probably due to sintering. c, which can be related to the energy of adsorption,[18] rose sharply above about 300°C to values similar to those already reported for the adsorption of nitrogen on iron.[18] The low values for lower outgassing temperatures were probably due to the presence of adsorbed layers of gases, which were removed only by more severe outgassing. The important point, however, is that there was no abrupt change in the surface area corresponding to this sharp rise in c. The presence of these adsorbed layers apparently did not appreciably affect the surface area measurements, and use of the lower outgassing temperatures to avoid sintering was justified.

The precision of these measurements was of the order of 5 percent, but the absolute accuracy in terms of particle size was not as good. Sintering, purity of the gases used, completeness of outgassing, the presence of low pressure adsorption, the choice of the area occupied by each nitrogen molecule, surface roughness, and the contact areas between particles are all factors that could introduce errors. The experimental conditions were chosen to minimize the first three of these. Beeck[19] has shown that nitrogen may be quasi-chemisorbed on iron to the extent of half a monolayer at 77°K and at pressures less than 0.1 mm of mercury. If this occurred during the measurement of the adsorption

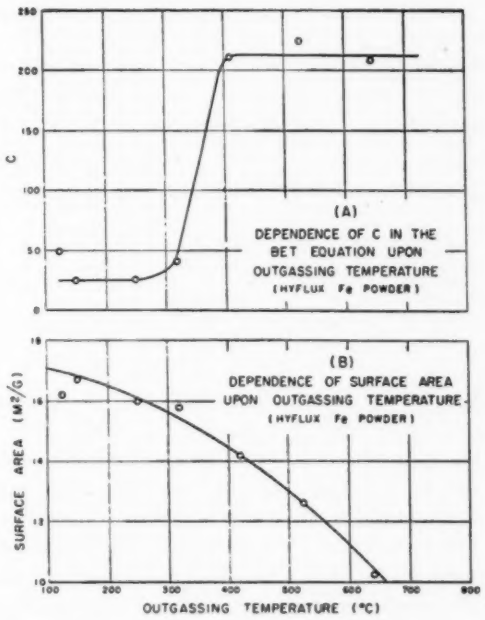


FIG. 3. Effect of outgassing temperature on the constants in the BET adsorption isotherm for N_2 on Fe powder.

isotherm, the measured surface area would be too large by a factor of 1.5. These results were true for Beeck's very clean vacuum-evaporated films, where the heat of adsorption was in the range from 10 to 5 kcal/mole. In the present case, however, with only mild outgassing, the observed c values indicate a heat of adsorption on these undoubtedly "dirty" surfaces of ≤ 2 kcal/mole. Much less tendency to chemisorption must have been present, and the uncertainty in the surface areas was probably less than indicated by the 1.5 factor.

Two values, $15.4A^2$ [4,6,20] and $16.2A^2$ [20–23] have been advanced for the area occupied per nitrogen molecule. The choice of which value to use is somewhat arbitrary, and introduces an uncertainty of about 5 percent in the results. A value of $15.4A^2$ was used in this case.

The packing of the particles produces areas of contact that are inaccessible to the nitrogen molecules. For the extreme example of closest packing of spheres 300A in diameter, about 10 percent of the surface area would be lost. The actual error was probably less than this, but was still quite uncertain. Schreiner[24] has reported that the amount of nitrogen adsorbed may be varied by as much as 10 percent through variations in packing of a powdered adsorbent, all other conditions being held constant.

The effect of surface roughness would be to increase the surface area, but it could not be estimated, beyond

[15] Brunauer, Emmett, and Teller, J. Am. Chem. Soc. **62**, 1723 (1940).
[16] P. H. Emmett, in "Symposium on New Methods for Particle Size Determination in the Subsieve Range," Am. Soc. Testing Materials (1941).
[17] A. C. Zettlemoyer and W. C. Walker, Ind. Eng. Chem. **39**, 69 (1947).
[18] S. Brunauer, *Physical Adsorption* (Princeton University Press, Princeton, 1943).
[19] O. Beeck, in *Advances in Catalysis* (Academic Press, Inc., New York, 1950), Vol. II, p. 155.
[20] H. K. Livingstone, J. Colloid Sci. **4**, 447 (1949).
[21] M. L. Corrin, J. Am. Chem. Soc. **73**, 4061 (1951).
[22] W. D. Harkins and G. Jura, J. Chem. Phys. **11**, 431 (1943).
[23] H. L. Pickering and H. C. Eckstrom, J. Am. Chem. Soc. **74**, 4775 (1952).
[24] H. Schreiner, Kolloid-Z. **123**, 113 (1951).

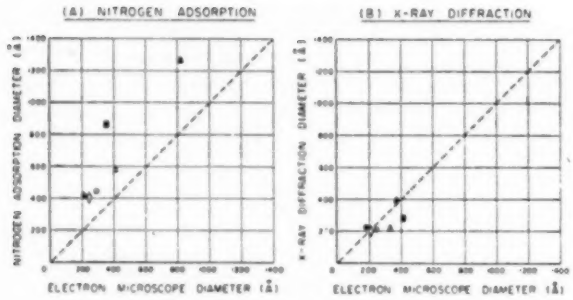


FIG. 4. Comparison of x-ray, N_2 adsorption, and electron microscope particle sizes.

the observation that in the micrographs the particles appeared to be relatively smooth.

There is a tendency for these uncertainties to oppose each other. Sintering, incomplete outgassing, and contact areas would make the observed surface area lower than the true value; low pressure adsorption, surface roughness, and a shift from 15.4 to 16.2A² per nitrogen molecule would all increase the observed surface area.

If an observed uncertainty of about 7 percent in the density measurements be included, the over-all uncertainty in the particle size was probably not better than about 20 percent.

In order to compare the particle sizes obtained from the surface areas with those obtained from the electron microscope, the average value expected from the former measurement may be calculated using the particle size distribution given by the latter. This average value, assuming spherical particles, is given by

$$d_A=\frac{\sum n_i d_i^3}{\sum n_i d_i^2}. \tag{6}$$

4. Discussion of Results

The observed data are collected in Table II. The measured diameters were obtained from the two subsidiary techniques, d_β and d_s from the x-ray measurements according to Eqs. (2) and (4), and d_A from the surface areas. The calculated diameters are the appropriate averages computed from the electron microscope particle size distributions, using Eqs. (3), (5), and (6) for d_β, d_s, and d_A, respectively.

The data for P-818 are included in Table II to illustrate the fact that the x-ray methods give crystallite rather than particle sizes. Only when each particle is a single crystal will these sizes be the same. In the present case, the agreement between the x-ray methods and the electron microscope indicates that this was probably true of all powders save the P-818. P-818 is a carbonyl iron, with particles which appeared in the microscope as regular spheres about 2 microns in diameter. The x-ray diffraction lines, however, were very broad, and indicated crystallite diameters on the order of 80A. After annealing the powder in hydrogen for two hours at 450°C, the x-ray lines were very much sharper, and were, in fact, almost indistinguishable from those obtained with well-annealed HP powder or with Armco iron. Apparently either small crystals within the particles had grown, or else internal stresses had been relieved. Taylor[25] has suggested that both small crystals and internal stresses are present in carbonyl iron and act to broaden the x-ray diffraction lines. Since the apparatus used here could not distinguish between these effects, all broadening has been treated as though due to small crystals.

The agreement between the x-ray methods and the electron microscope is in most cases satisfactory. The fact that both x-ray sizes agree fairly well with their electron microscope counterparts indicates an internal consistency in the x-ray method. In every case, the sizes derived from the surface areas are higher than their electron microscope counterparts. These relations are shown graphically in Fig. 4, in which the abscissas are the expected values from columns 3 and 5, and the ordinates the measured values from columns 2 and 4 of Table II. If perfect agreement were obtained between the electron microscope and the other two methods, the data would fall on the 45° lines.

One reason for the larger sizes found with the surface area measurements may be the presence of appreciable amounts of oxide in the powders. Since the x-ray methods make use of a diffraction line in the iron pattern, they probably recorded the crystal size of the metallic portion only of the particles. Since the electron microscope agrees with the x-ray methods, it also apparently gave only the size of the metallic part, due perhaps to the relative transparency of the oxide. On the other hand, if the oxide were present as surface layers, the surface area measured would be that of the whole particle, and the particle size would be larger than that given by the other methods. There is, however, no apparent correlation between the percentages of free metal in the samples and the deviations of the surface area sizes as compared to the others.

These deviations agree qualitatively with previous observations. Birks and Friedman[7] found that the electron microscope and x-ray sizes of MgO particles agreed to within about 10 percent. Zettlemoyer and Walker[6] compared x-ray and surface area sizes for MgO, and

TABLE II. Comparison of particle sizes.

Material	d_β Meas.	d_β Calc.	d_s Meas.	d_s Calc.	d_A Meas.	d_A Calc.
P-818, Unannealed	88A		82A		1.8μ	1.6μ
P-818, Annealed						
Fe-8	389A	370A	370A	224A	1270A	820A
Fe-7	282	406	212	233	862	350
Hyflux	213	329	147	272	596	407
NOL-1	207	254	140	201	444	298
NOL-2	228	171	178	143	407	214
Ugine	200	206	150	168	404	254

[25] A. Taylor, Tenth Annual Pittsburgh Conference on X-Ray and Electron Diffraction (1952).

found in every case that the x-ray size was smaller. Kopelman[9] also found that x-ray sizes were invariably smaller than those determined from nitrogen adsorption isotherms, for fine iron powders.

These investigators calculated the x-ray sizes using an approximate correction for instrumental broadening introduced by Warren.[26] In Table III a comparison is made among x-ray sizes calculated using the Stokes Fourier transform technique, Warren's approximation, and an approximate correction procedure suggested by Jones,[13] which was designated by him as procedure *b*. In all three computational methods the integral breadth of the pure diffraction line is first calculated, and then the Scherrer equation, Eq. (2), is used to arrive at the particle size. The differences apparent in Table III are therefore purely computational, and since the Stokes method involves no assumptions concerning the apparatus, as the others do, the differences must be considered as errors on the part of the other methods. Alexander and Klug[14] have pointed out similar errors in the approximate methods with other materials. With the Jones method, in particular, the errors are not large. Considering the simplicity to which the Patterson-Tunnell strips reduce the Stokes calculation, however, use of the latter seems definitely worthwhile.

The conclusions drawn from this work can be summarized as follows.

(1) Reasonably good electron microscope size distributions may be obtained with submicron particles of ferromagnetic metals, if care be taken to obtain good dispersions. For single-domain particles, this may be done with the ac magnetic field method described above.

[26] B. E. Warren, J. Appl. Phys. **12**, 375 (1941).

TABLE III. Comparison of x-ray sizes.

Material	Stokes, d_β	Warren	Jones
Fe-8	389A	325A	425A
Fe-7	282	250	307
Hyflux	213	192	222
NOL-1	207	185	205
NOL-2	228	168	220
Ugine	200	165	183

(2) In order to have a check upon these distributions, independent measurements of at least one average size should be made. The x-ray technique apparently is to be preferred. It must be remembered, however, that each method measures something different, and perfect agreement among these three cannot always be expected. An extreme example in which agreement should not be expected would be a material containing polycrystalline particles with very irregular surfaces.

(3) With the x-ray technique, the Stokes Fourier transform method should be used. Not only is it a rigorous computational method, but it also makes possible the calculation of at least two different average sizes. Under favorable experimental conditions the entire particle size distribution may be obtained by this method.[2,3] It can, therefore, provide an excellent check upon the electron microscope results.

ACKNOWLEDGMENT

Valuable contributions to this work were made by L. Muldawer with the interpretation of the x-ray data, by F. Jaumot with both the x-ray and electron microscope phases, and by R. Smith and D. Husted, with some of the electron micrographs. The support of the U. S. Office of Naval Research is gratefully acknowledged.

Determination of Reflection Coefficients and Insertion Loss of a Wave-Guide Junction*

GEORGES A. DESCHAMPS
Federal Telecommunication Laboratories, Inc., Nutley, New Jersey
(Received April 2, 1953)†

Reflection and transmission coefficients are important parameters of a wave-guide junction. It is shown that these parameters in phase as well as in magnitude may be read directly from a graph. A short-circuit is moved in one of two wave guides that are connected by the junction, and the corresponding reflection coefficient is measured in the other wave guide as in a standard method. The positions of the short-circuit, however, are taken equally spaced every $\frac{1}{8}$th of a wavelength. When plotted in the complex plane, the corresponding reflection coefficients fall on a circle and the chords joining opposite points intersect at a point that is simply related to the desired parameters.

When very accurate measurements are desired, 8 or even 16 short-circuit positions spaced at $\frac{1}{16}$th or $\frac{1}{32}$th wavelength can be used. The quality of the chord intersection then gives a control of the experiment by showing systematic or random errors. When the intersection is not perfect, because of random errors, a method of averaging is described that reduces these errors.

I. INTRODUCTION

THE best description of a wave-guide junction from the viewpoint of power relations is the scattering matrix.[1] For the elements of this matrix, s_{11} is the reflection coefficient at terminal 1 when line 2 is matched, s_{22} is the reflection coefficient at terminal 2 when line 1 is matched, and s_{12} is the transmission coefficient from 1 to 2. Terminals 1 and 2 are some arbitrarily specified points in the two transmission lines (Fig. 1), and it is understood that the two wave guides or transmission lines can carry only one mode. The transmission coefficient s_{21} is equal to s_{12} when reciprocity applies, as will be assumed here. The insertion loss

$$(s_{12})_{\text{db}} = -20 \log_{10} |s_{12}| \tag{1}$$

is a measure of the efficiency of the junction.

The junction is generally described by some equivalent circuit, which is obtained by impedance measurements and from which the scattering coefficients would have to be deduced. The present method gives them more directly by interpreting the bilinear relation

$$w' = T(w) \tag{2}$$

between a reflection coefficient w at terminal 2 and its image w' at terminal 1.

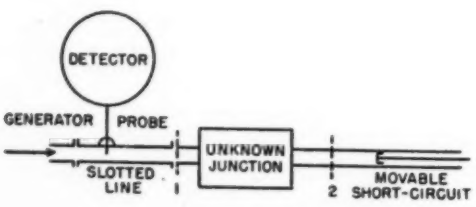


FIG. 1. Experimental setup.

II. MEASUREMENT PROCEDURE

The experimental procedure is a slight modification of the conventional method.[2] A short-circuit is moved in line 2 and produces at terminal 2 a reflection coefficient of unit amplitude and known phase. Measuring the standing waves in line 1 will give the corresponding reflection coefficient w' at terminal 1. (See Fig. 1.)

Instead of moving the short-circuit in line 2 arbitrarily, we displace it in equal steps, for instance, $\frac{1}{16}$th-wavelength intervals. Thus the reflection coefficient w takes values $w_1, w_2, \cdots, w_8$, which form a regular polygon on the unit circle Γ in the w plane [Fig. 2(a)]. The images by the transformation T, characteristic of the junction, fall on a circle Γ' inside Γ as shown in Fig. 2(b).

The diameters of the unit circle Γ become circles orthogonal to Γ' and passing through the point O', the image of O. The point O', called the iconocenter[3] of Γ', plays an important part in the interpretation of the results.

III. REFLECTION COEFFICIENT s_{11}

Determination of the reflection coefficient s_{11} is equivalent to the construction of the iconocenter, since obviously the complex number corresponding to O' is

$$T(O) = s_{11}. \tag{3}$$

The point O' may be obtained by drawing the circles shown in Fig. 2(b). However, it is simpler to proceed as follows. The chords $w_1'w_5'$, $w_2'w_6'$, $w_3'w_7'$, $w_4'w_8'$, have a common crossover point $\bar{O}$. If as in Fig. 3 the perpendiculars CL and $\bar{O}K$ are drawn to the radius $C\bar{O}$ of circle Γ', the point O' will be the intersection of $C\bar{O}$ and LK.

The existence of the crossover point and its relation to the iconocenter may be proved by considering the sphere S having Γ' as its equator (Fig. 4). By stereo-

* This work was sponsored by the Wright-Patterson Air Force Base, Ohio.
† The original unrevised manuscript was received September 29, 1952.
[1] Montgomery, Dicke, and Purcell, *Principles of Microwave Circuits* (McGraw-Hill Book Company, Inc., New York, 1948), p. 146.
[2] N. Marcuvitz, *Waveguide Handbook* (McGraw-Hill Book Company, Inc., New York, 1951), pp. 130–138.
[3] From the Greek εικων for image. The point O' is not defined uniquely by the circle Γ' but depends on how the images w_i' are distributed on it.

graphic projection from the pole L, any circle γ' passing through O' and orthogonal to Γ' is transformed into a circle γ on S, also orthogonal to Γ' and passing through the stereographic projection K of O'. By projection on the plane of the equator, this circle becomes a straight line $\bar{\gamma}$, which goes through the projection $\bar{O}$ of K. The construction in Fig. 3 is a reproduction of $LCO'K\bar{O}$ on the plane of Γ' obtained for instance by rotation through 90 degrees about CO'.

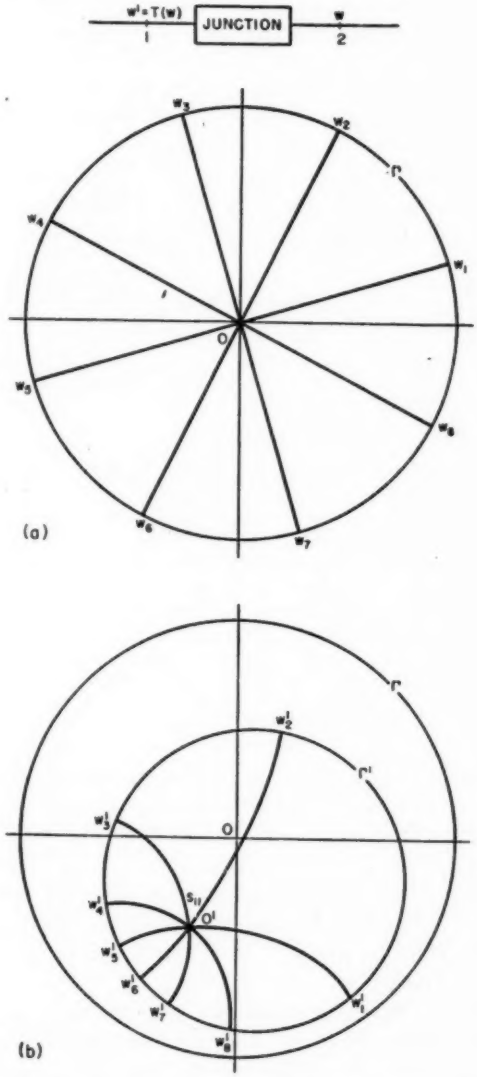


FIG. 2. Unit circle Γ and its image Γ' by the transformation T, characteristic of the junction. (a) Reflection coefficient w at terminal 2. (b) Reflection coefficient w' at terminal 1.

It is to be noted that s_{11} is obtained correctly by the present method even if the short-circuit is not perfect. The circle Γ is then replaced by a circle concentric to it and therefore its image has the same iconocenter O' as before.

If a matched load is available in line 2, the point O' can also be obtained by direct measurement of the corresponding input reflection coefficient in line 1.

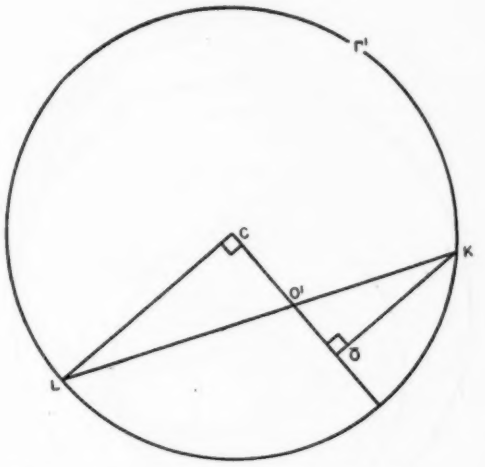


FIG. 3. Relation between the iconocenter and the crossover point.

IV. POWER RELATIONS

The power relations that are expressed by the absolute values of the scattering coefficients can be deduced from the iconocenter O', the center C, and the radius R of the image circle Γ'. The equations proved in Sec. VI are

$$|s_{11}| = OO', \tag{4}$$

$$|s_{22}| = CO'/R, \tag{5}$$

$$|s_{12}|^2 = R(1-|s_{22}|^2). \tag{6}$$

An alternative form of Eq. (6) deduced by elementary geometry is

$$|s_{12}| = O'H/R^2, \tag{7}$$

where $O'H$ is perpendicular to CO' and H is on Γ' (Fig. 5). Introducing the notation $(x)_{db}$ to mean

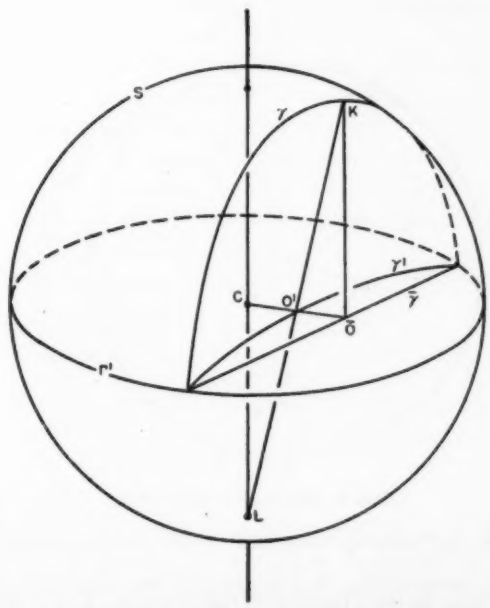


FIG. 4. Transformation from the crossover point to the iconocenter.

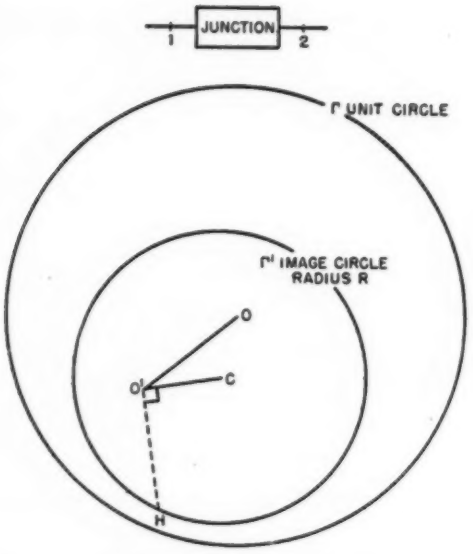


FIG. 5. Power relations. The magnitudes of the reflection coefficient on sides 1 and 2 of the junction are, respectively, $|s_{11}| = OO'$ and $|s_{22}| = CO'/R$. The magnitude of the transmission coefficient is $|s_{12}| = O'H/R$ or in decibels the insertion loss is $(s_{12})_{db} = (O'H)_{db} - \frac{1}{2}(R)_{db}$.

$-20 \log_{10}|x|$, we find the insertion loss

$$(s_{12})_{db} = (O'H)_{db} - \tfrac{1}{2}(R)_{db}. \tag{8}$$

These results do not depend on the choice of the terminal planes for the junction.[4]

When a wave of unit amplitude is incident on side 2 of the junction and the transmission line connected to 1 is matched, the power $|s_{22}|^2$ is reflected and $|s_{12}|^2$ is transmitted. Therefore, R indicates, according to Eq. (6), the fraction of the net incident power $1 - |s_{22}|^2$ that is transmitted.[5]

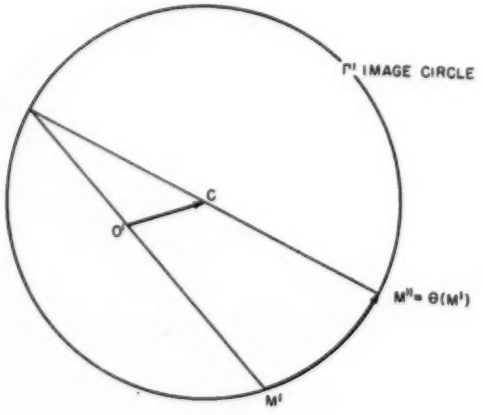


FIG. 6. The transformation θ.

[4] They are "distance invariant" in the sense of reference 2, p. 137.

[5] This interpretation of R is found in W. Altar, Proc. Inst. Radio Engrs. 35, 355 (1947); 478 (1947); and also in A. L. Cullen, Wireless Engr. 26, 255 (1949).

V. PHASE RELATIONS

Before considering the phases of s_{22} and s_{12}, it is convenient to introduce the transformation Θ, which is the result of a negative inversion of center O' leaving Γ' invariant followed by a negative inversion of center C also leaving Γ' invariant. The effect of Θ on the points of Γ' is quite simple; it is a projection of center O' followed by a projection of center C and transforms M' into M'', for instance, as illustrated in Fig. 6. Θ is a bilinear transformation that carries O' into C and, therefore, it also carries the circles orthogonal to Γ' and passing through O' into the diameters of Γ'. Since the transformation is conformal, these diameters form between them the same angles as the original diameters of Γ.

The use of Θ makes it possible to extend the transformation T to points of Γ for which a measurement has not been made. For instance, if the point P $(w=+1)$ is

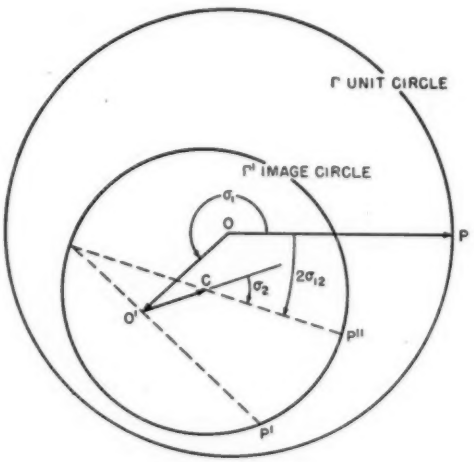


FIG. 7. Phase relations. The phases of the scattering coefficients s_{11}, s_{22}, s_{12} are, respectively, σ_1, σ_2, σ_{12}.

not one of the w_i, we find its image by considering $w_1'' = \Theta(w_1')$ and observing that the angle between P'' and w_1'' on Γ' must equal the angle between $+1$ and w_1 on Γ. We can construct P'' from this property and then P' by applying Θ^{-1}.

As shown in the next section, the phase σ_2 of s_{22} is then the angle between the vectors $\mathbf{O'C}$ and $\mathbf{CP''}$ (Fig. 7),

$$\sigma_2 = (\mathbf{O'C}, \mathbf{CP''}). \tag{9}$$

The phase[6] σ_{12} of the transmission coefficient s_{12} is

$$\sigma_{12} = \tfrac{1}{2}(\mathbf{OP}, \mathbf{CP''}). \tag{10}$$

It is to be noted that the sign of s_{12} cannot be determined by the present method and this is reflected in Eq. (9), which gives σ_{12} only modulo π.

[6] The angles in Eqs. (9) and (10) are directed angles between directed segments or vectors. Changing the sense of one of these vectors or their order would give an incorrect result.

VI. RELATIONS BETWEEN IMAGE CIRCLE AND SCATTERING COEFFICIENTS

The transformation T is given explicitly by

$$w'=T(w)=s_{11}+\frac{s_{12}^2 w}{1-s_{22}w}. \tag{11}$$

One verifies immediately that

$$T(O)=s_{11}.$$

To find the other relations, first assume that the terminals have been chosen at 1′ and 2′ (Fig. 8) so as to make both s_{11} and s_{22} real and positive.

The point C and the point at infinity are inverse with respect to Γ', and therefore $T^{-1}(C)$ and $T^{-1}(\infty)=s_{22}^{-1}$ are inverse with respect to the unit circle Γ. The inverse of the point $1/s_{22}$ being the point $s_{22}^*=s_{22}$, the affix[7] of

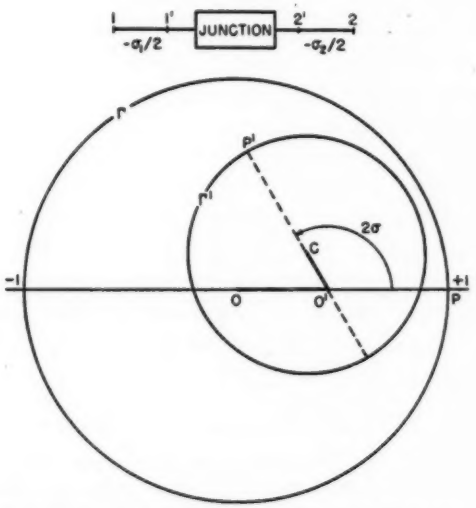


FIG. 8. Image circle when both reflection coefficients are real and positive. (Junction between terminals 1′ and 2′.)

C will be

$$w_C=T(s_{22})=s_{11}+\frac{s_{22}s_{12}^2}{1-s_{22}^2}. \tag{12}$$

The image P' of the point P ($w=+1$) corresponds to

$$w_1=T(1)=s_{11}+s_{12}^2/(1-s_{22}). \tag{13}$$

The affix of $\mathbf{CP'}$ is, therefore,

$$w_1-w_C=s_{12}^2/(1-s_{22}^2), \tag{14}$$

where the phase angle is twice that of s_{12} and where the absolute value is the radius R of the image circle Γ',

$$R=\frac{|s_{12}|^2}{1-s_{22}^2}. \tag{15}$$

[7] The affix of a point or of a vector in the complex plane is the complex number corresponding to that point or that vector.

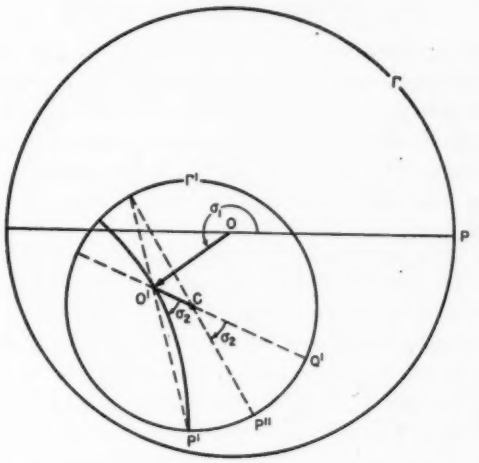


FIG. 9. Transformation of the image circle (Fig. 7) when the junction is referred to its actual terminals 1 and 2.

The affix of $O'C$ is

$$w_c-s_{11}=\frac{s_{22}s_{12}^2}{1-s_{22}^2}.$$

It has the same phase 2σ as w_1-w_c, and therefore the points $O'CP'$ are on a straight line as in Fig. 8 and also

$$s_{22}=O'C/R. \tag{16}$$

It remains to show what happens when s_{11} and s_{22} have the phases σ_1 and σ_2, respectively. The junction is now considered between its terminals 1 and 2 instead of 1′ and 2′ and the lengths of line 11′ and 22′ are, respectively, $-(\sigma_1/2)$ and $-(\sigma_2/2)$ radians.

The effect of the line on side 1 is to rotate the circle Γ' and the points on it about the point O through the angle σ_1 (Fig. 9). The effect on side 2 is to rotate the points on the unit circle Γ through the angle σ_2 before they undergo the above transformation. The image of P ($w=+1$) instead of being the point Q' on $O'C$ produced, as it would be if $\sigma_2=O$, is the point P'. Therefore σ_2 is the angle between $O'C$ and the arc of a circle $O'P'$, image of OP. Applying the conformal transformation Θ considered in Sec. V, this angle becomes $(\mathbf{CQ'}, \mathbf{CP''})$ or $(\mathbf{O'C}, \mathbf{CP''})$ as indicated in Eq. (9).

For the phase σ_{12} of s_{12}, we have to add $\sigma_1/2$, $\sigma_2/2$, and σ:

$$\begin{aligned}2\sigma_{12}&=\sigma_1+\sigma_2+2\sigma\\&=(\mathbf{OP}, \mathbf{OO'})+(\mathbf{O'C}, \mathbf{CP''})+(\mathbf{OO'}, \mathbf{O'C})\\&=(\mathbf{OP}, \mathbf{CP''}),\end{aligned} \tag{17}$$

which justifies Eq. (10).

Equations (6) and (5) result from (15) and (16) by replacing s_{22} by its absolute value if we note that the length $O'C$ and R are not changed by the introduction of the lines 11′ and 22′.

VII. APPROXIMATE RELATIONS

When a rapid evaluation is desired, it is sufficient to take four positions of the short-circuit at intervals $\frac{1}{8}$th wavelength apart. Three positions would be sufficient and, even if taken arbitrarily, could be used to construct the iconocenter. However the simplification in interpreting the measurement justifies taking an extra reading.

The following remarks help in computing the scattering coefficients. If the crossover point is not more than $R/4$ away from the center C of the circle Γ', the phase angles on Γ' can be obtained with less than 1 degree of error by measuring them directly from the crossover point O. The iconocenter is then practically the middle of $C\bar{O}$ with less than 1.5 percent error on CO'. Finally when s_{22} is small, Eq. (5) reduces to $|s_{12}| \simeq R^{\frac{1}{2}}$.

VIII. PRECISION MEASUREMENT

If on the other hand very accurate results are desired, it becomes advisable to use 8 or even 16 points equally spaced on the circle Γ. Then, two controls on the quality of the experiments are available: The image points should fall on a circle and the 4 (or 8) chords should intersect at a single crossover point. When these two tests do not succeed exactly and the obvious reasons for systematic errors[8] have been checked, the residual random errors can be minimized by a double-averaging procedure. The first one, which applies to the conventional method as well, consists of passing the best circle through the w_i' points. The other is based on estimating the crossover point, then applying the transformation Θ and measuring the phases of the points w_i'' on Γ'. These should be equal to the phases of the original w_i. The plot of one phase *versus* the other should be a straight line with a 45-degree slope. If no systematic trend is observed, which would indicate a bad choice of crossover point, the fitting of a straight line (by the least-square method, for instance) gives a good representative point (P' image of P for instance) for the set of experimental data.

[8] Some of the errors are detected by the two tests. Accidental errors show up as stray points. Strong attenuation in line 2 produces a circle that does not close or rather an arc of a spiral. Nonuniform lines produce bad crossover. An error in the wavelength gives chords tangent to a small ellipse instead of intersecting.

IX. CONCLUSIONS

The method of analysis presented here differs from the conventional one[2,9,10] in the use it makes of the strong relation that exists between the points on the circle and their images by a bilinear transformation. In a disguised way, it uses the invariance of the cross-ratio.[11] This affords a new control on the measurement and a means of averaging by fitting a *straight line* to the experimental results properly treated.

In the conventional method applied to lossless structures, one would fit to these results a tangent relation while for lossy structures the procedure is more complicated. In the present method, the consideration of lossy structures does not introduce any difficulties.

For ordinary measurements, the experiment is practically the same but its interpretation is reduced to a minimum: drawing of a few straight lines and direct reading of angles and distances. It gives immediately the most important parameters of the junction transmission and reflection coefficients, independently of any equivalent circuit representation and of a more-or-less-arbitrary convention for the characteristic impedances of the wave guides.

[9] A. Weissfloch, Hochfrequenztechnik und Elektroakustic **60**, 67 (1942).

[10] N. Marcuvitz, Proc. Inst. Radio Engrs. **36**, 728 (1948).

[11] It follows that the same method is applicable to find the analytic expression of any bilinear transformation even when the scattering interpretation does not apply. This covers problems on interference, polarization transformers, etc.

Erratum: Determination of the Drag on a Cylinder at Low Reynolds Numbers

[J. Appl. Phys. **24**, 771–773 (1953)]

R. K. FINN
University of Illinois, Urbana, Illinois

EQUATION (1) should read as follows:

$$C_D = \frac{8\pi}{N_{Re}[\frac{1}{2} - \gamma - \ln(N_{Re}/8)]}$$

A Replica Technique in the Study of Chemical Precipitation Processes*

ROBERT B. FISCHER AND JOSEPH P. ELLINGER
Department of Chemistry, Indiana University, Bloomington, Indiana
(Received April 13, 1953)

A replica technique has been developed for the electron microscope study of particles of precipitates formed by mixing solutions of chemical reagents. The replica is formed thermally in a thin film of parlodion. A one-step, shadowed replica has been found most satisfactory. The technique has been found suitable for studies of the effect of conditions of observation upon the specimen materials, studies of precipitates such as silver chloride which are known to be unstable during direct observation, studies of precipitates as they exist in the mother liquid; and studies which require detailed observation of surfaces.

INTRODUCTION

THE electron microscope has been employed in a number of studies of precipitation processes as they are encountered in analytical chemistry. Although the direct observation in the electron microscope of precipitated particles has been of much value in this type of study, there are there reasons why it has appeared that an indirect replica technique might be of value. First, a number of precipitates which are of analytical interest are decomposed and/or contaminated through the combined influences of the evacuated column and irradiation with the electron beam. Second, many effects which are of interest are predominately surface effects, so that a method of surface examination should be of value. Third, a technique permitting (indirect) observation of the precipitated particles while still suspended in mother liquid should be useful. Therefore, an attempt was made to find or to develop a replica technique suitable for the examination of particles of precipitates formed by mixing solutions of chemical reagents.

Most of the conventional replica methods are not suitable for the replication of particles of this type. However, a replica technique was developed by Fischbein[1] for the wet replication of biological specimens, and this technique served as a starting point in the present investigation.

DESCRIPTION OF TECHNIQUE

The replica is obtained thermally in a sandwich consisting of a glass slide, thin film of replicating plastic, drop of the mother liquid containing the suspended precipitate, and another glass slide. The sandwich is tightened together with a screw-type pinch clamp and then immersed in the mother liquid and heated to boiling for five or ten minutes. The sandwich is then removed from the liquid, the clamp removed, and the plastic film allowed to dry on the lower glass slide. The precipitate is then dissolved with a suitable solvent and the replica washed with a few drops of distilled water. The plastic film is then floated off on a water surface and portions mounted on nickel grids conventionally. As is usually the case with thin plastic replicas, shadowing is almost necessary to provide adequate contrast in the final image.

The several possible variables in the procedure were investigated as follows.

1. Choice of replicating material. This material must meet the usual requirements for specimen mounting membranes, and in addition it must soften below the boiling point of water. Although other materials are doubtless satisfactory also, parlodion was found to serve very well.

2. Thickness of film. Unless the initial replica is to be replicated in turn, the film must be of a thickness reasonably transparent to the electron beam. Although the exact thickness is not critical and although other combinations of conditions are quite acceptable, good replicas were obtained in films formed by dipping a clean microscope slide into a 2 to 3 percent solution of parlodion in amyl acetate and drying it in the vertical position.

3. One-step or multi-step replica process. Although Fischbein[1] recommended a three-step process with wet biological specimens and a final replica of silica, most satisfactory results were obtained in the present study with a simple one-step replica process. A second replica of silica appeared to yield pictures that looked as good to the eye, but the procedure was greatly lengthened with no evident advantage. Therefore, the one-step process is to be recommended.

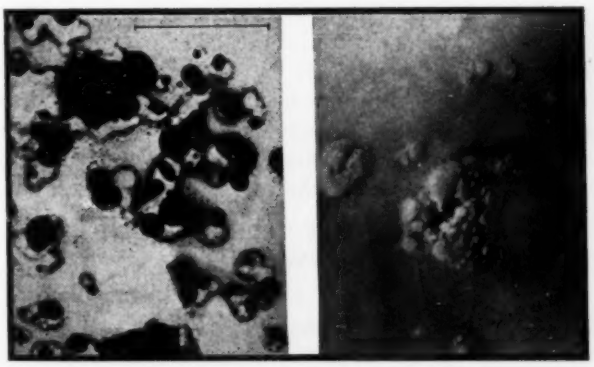


FIG. 1. Micrographs of silver chloride. Direct view (left). Replica (right)

* Publication number 583 from the Chemistry Department of Indiana University, Bloomington, Indiana.

[1] I. W. Fischbein, J. Appl. Phys. 21, 1199 (1950).

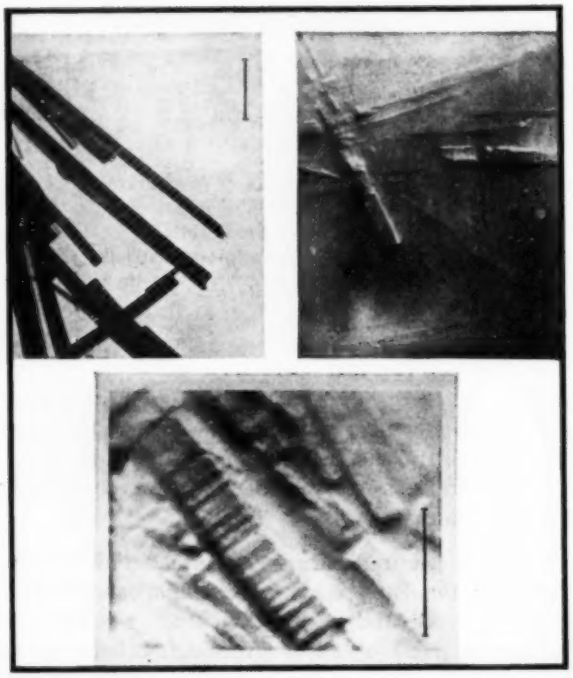


FIG. 2. Micrographs of nickel dimethylglyoxime. Direct view (upper left) Replicas (upper right and lower)

4. Type of shadowing material and thickness of deposit. Shadowing is essential to obtain images of adequate contrast. Several of the common shadowing materials were tried with no apparent unusual advantage of one over another. Conventional considerations also apply with regard to thickness of the shadowing deposit.

5. Angle of shadowing. The choice of shadowing angle involves considerations somewhat different than in shadowing conventional particulate specimens because the shadow appears entirely inside the image of the particle rather than outside. Both regions of the shadow, i.e., the regions containing excess metal and those containing deficiency of metal, are actually inside the image of the particle in the negative replica. If the shadowing angle is too small, the shadowing metal may even miss the valley caused by the specimen particle, whereas with a shadowing angle approaching the normal to the surface the edge patterns in the shadowed specimen would be much narrower. These factors are of significance in interpreting the resultant micrographs. Various shadowing angles were used in the preliminary work, while an angle of thirty degrees was used with most of the more recent preparations.

ILLUSTRATIONS OF TECHNIQUE AND ITS APPLICABILITY

The micrographs of Fig. 1 are of silver chloride, a typical curdy precipitate. The left-side picture is a direct micrograph of silver chloride particles, which are partially decomposed at the time of observation. The right-side picture is a replica of silver chloride particles formed while the particles were still immersed in the mother liquid from which they were precipitated. This particular replica has been chosen for this illustration, because it emphasizes a point of significance in interpretation of this type of replica. The opaque material is still-undissolved silver chloride; since its shadow is opposite in direction to the shadow of the replicated areas, and since the silver chloride must represent an elevation on the surface, the replicated areas are truly depressions on the surface of the film, appearances to the contrary notwithstanding.

Crystals of nickel dimethylglyoxime are seen in the upper left picture of Fig. 2, while replicas of the same material are shown in the upper right picture. The general nature of the replica images is evident in this micrograph of a replica of the rod-like crystals of nickel dimethylglyoxime. A replica view of one of these particles is shown at higher magnification in the lower picture of Fig. 2. The striated structure on the surface has been seen repeatedly on replicas of this substance,

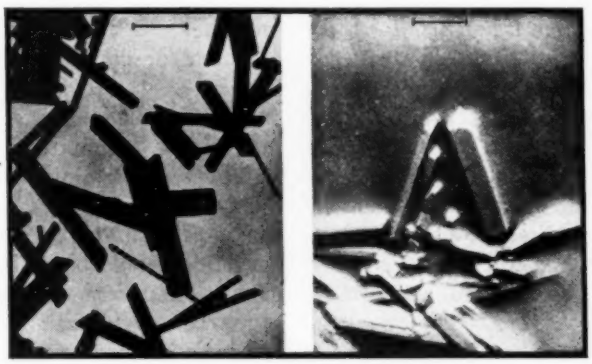


FIG. 3. Micrographs of lead chromate. Direct view (left). Replica (right)

whereas it is not at all discernible on the direct electron micrographs of the crystals.

In Fig. 3 are shown a direct micrograph of crystals of lead chromate along with a micrograph of a replica of the same material formed while the crystals were still suspended in mother liquid.

EVALUATION OF METHOD

A method for the replication of chemically precipitated substances has been developed and found readily usable. It has been found useful in studies of the following types: studies of the effect of conditions of observation in the electron microscope upon the specimen materials; studies of precipitates such as silver chloride which are known to be unstable during direct observation; studies of precipitates as they exist in the mother liquid; studies which require detailed observation of the surfaces.

Among the possibly less desirable features of the method are the following: the procedure is somewhat more time consuming than is direct viewing of the

precipitated particles; the interpretation of the micrographs is somewhat different than that of conventionally shadowed particulate specimens; any shrinkage of the film, especially if not uniform, would cause apparent distortion in the appearance of the specimen in this as in all other plastic replica studies; the pressure required at the time of replication may fracture relatively delicate crystals. In addition, the fact that surface structure alone is seen in the replica may occasionally be undesirable, although this is more often than not a strong feature in the applicability of this technique.

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 8 AUGUST, 1953

Methods of Measuring the Properties of Ionized Gases at High Frequencies. IV.* A Null Method of Measuring the Discharge Admittance

LAWRENCE GOULD AND SANBORN C. BROWN
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received April 13, 1953)

The admittance of a gas discharge may be obtained from measuring the ratio of the transmitted power through a microwave cavity to the incident power as a function of signal frequency near the cavity resonance. The method involves balancing the transmitted and incident signals to zero at the cavity resonance after they have passed through two separate receiving systems. When the signal frequency is changed, an attenuator is used to rebalance the two signals. The change in frequency from resonance and the corresponding change in attenuation gives the necessary data which will plot as a straight line whose slope yields the desired information.

I. INTRODUCTION

IN three previous papers referred to as Parts I, II, and III,[1] methods for measuring the complex admittance and electron density of a discharge are discussed. Resonance curves obtained by means of standing wave measurements give the data necessary for determining these quantities. A simpler and more precise technique for obtaining the discharge characteristics will be considered here. The method consists of measuring the ratio of the power transmitted through a microwave cavity to the power incident as a function of signal frequency in the vicinity of the cavity resonance. By proper choice of variables, the data will plot as a straight line whose slope yields the desired information. This reduces the data necessary and enhances the accuracy of the measurements. This method may be used to measure the characteristics of a discharge either in the steady state or transient condition.

II. THEORY

A microwave cavity containing a discharge may be represented by the equivalent circuit in Fig. 1. A detailed discussion of this general equivalence and the significance of the various quantities are given in Part I. The discharge admittance is obtained from the knowledge of g_d and b_d, the discharge conductance and susceptance. The empty cavity conductance and susceptance are presented by g and b, and g_s represents the conductance of the line and of the input coupling. Since the present technique involves measuring the transmitted power, the influence of the output loop and line has to be considered. If a very small output coupling is utilized, it may be represented by g_n, which is the conductance of the output line reflected back into the cavity. The power transmitted through the cavity is represented by the power absorbed in g_n and is designated by P_t. The object of this analysis is to find an expression for P_t as a function of the equivalent circuit parameters and the incident power P_i.

The normalized impedance of the cavity and discharge terminating the line at PP' in Fig. 1 is

$$z=1/g_s+1/(g_t+jb_t), \tag{1}$$

where $g_t=g+g_n+g_d$ and $b_t=b+b_d$. The power absorbed is $P_a=P_i(1-|\Gamma|^2)$ where $|\Gamma|$ is the magnitude of the complex reflection coefficient. The complex reflection coefficient is related to the impedance by the expression

$$\Gamma=(z-1)/(z+1). \tag{2}$$

The absorbed power is obtained by combining Eqs. (1) and (2)

$$\frac{P_a}{P_i}=\frac{4[(g_t^2+b^2)/g_s+g_t]}{[(1/g_s+1)g_t+1]^2+b^2(1/g_s+1)^2]}. \tag{3}$$

Equation 3 represents the power that is dissipated in g_s and g_t, whereas the power dissipated in g_n is the desired quantity. The power transmitted is related to

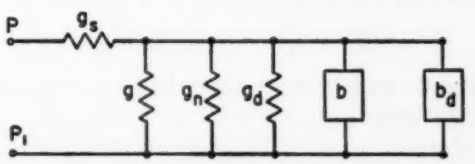


FIG. 1. Equivalent circuit for a resonant cavity.

* This work has been supported in part by the Signal Corps, Air Materiel Command, and U. S. Office of Naval Research.
[1] S. C. Brown and D. J. Rose, J. Appl. Phys. 23, 711 (1952); 719 (1952); 1028 (1952).

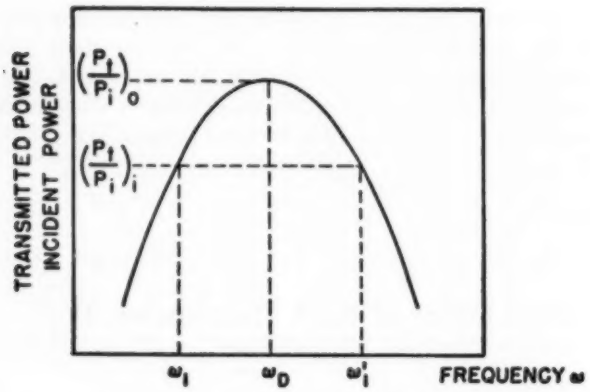


FIG. 2. Plot of the ratio of transmitted to incident power as a function of frequency.

the power absorbed through the expression

$$P_t = P_a(g_n/g_t)\frac{[g_t/(g_t^2+b^2)]}{[g_t/(g_t^2+b^2)+1/g_s]}, \quad (4)$$

where the factor

$$\frac{g_t/(g_t^2+b^2)}{g_t/(g_t^2+b^2)+1/g_s}$$

represents the fraction of the absorbed power dissipated in g_t, and g_n/g_t represents the fraction of power dissipated in g_n.

Combining Eqs. (3) and (4) yields

$$P_t = P_i 4g_n\{[(1/g_s+1)g_t+1]^2+b^2(1/g_s+1)^2\}, \quad (5)$$

where

$$b_t = \beta(\omega/\omega_0-\omega_0/\omega)+b_d. \quad (6)$$

The cavity resonance frequency with a discharge present, denoted by ω_0', is defined according to Eq. (4) of Part III as

$$\omega_0' = \omega_0 - b_d\omega/2\beta. \quad (7)$$

Since we are assuming that the presence of a discharge disturbs the cavity characteristics only slightly, that is $(\omega_0'-\omega_0)\ll\omega_0$, near the resonant frequency ω_0', Eq. (6) becomes

$$b_t = 2\beta(\omega-\omega_0')/\omega = \beta x/\omega, \quad (8)$$

where $x=2(\omega-\omega_0')$. Under these conditions, Eq. (5) takes the following form

$$\frac{P_i 4g_n}{P_t[(1/g_s+1)g_t+1]^2} = 1+\frac{\beta^2x^2(1/g_s+1)^2}{\omega^2[(1/g_s+1)g_t+1]^2} \equiv \gamma. \quad (9)$$

If γ is plotted as a function of x^2, the square root of the reciprocal of the slope of the line, represented by α is

$$\alpha = \omega[(1/g_s+1)g_t+1]/\beta(1/g_s+1). \quad (10)$$

The difference between the value of α with and without a discharge is

$$\Delta\alpha = \omega g_d/\beta \quad (11)$$

$$\Delta\alpha/2(\omega_0-\omega_0') = g_d/b_d. \quad (12)$$

Thus from a measurement of P_i/P_t as a function of frequency for a cavity with and without a discharge, it is possible to obtain the ratio of discharge conductance to discharge susceptance and the resonant frequency shift which is due to the presence of the discharge. In general, these measurements enable one to determine the complex conductivity of the ionized plasma σ and the electron density as shown in Part III. The ratio of the real part of the electron conductivity to the imaginary part may be obtained from Eq. (12) since according to Part III

$$\sigma_r/\sigma_i = g_d/b_d. \quad (13)$$

If the experimental situation is such that g_s may be neglected, then the unloaded Q of the cavity may be obtained directly from α provided the cavity is very much undercoupled at resonance. Under these conditions, $g_s > g_t$ and $g > 1$, and Eq. (10) becomes

$$\alpha = \omega g_t/\beta. \quad (14)$$

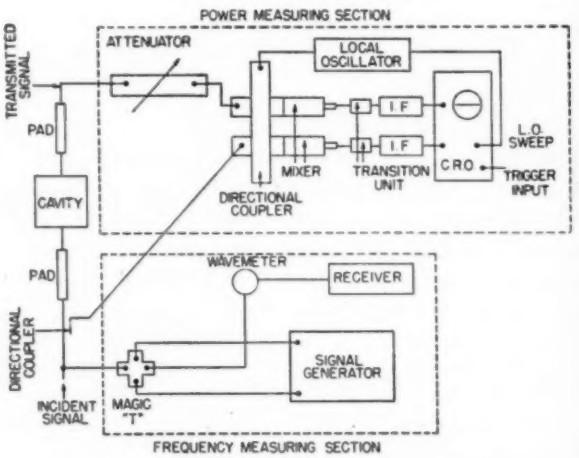


FIG. 3. Black diagram of measuring equipment.

According to Eq. (21a) Part I, Q_u is given by the expression $Q_u=\beta/g_t$, and hence we may rewrite Eq. (15) as

$$\alpha = \omega/Q_u \quad \text{or} \quad Q_u = \omega/\alpha. \quad (15)$$

In practice the condition that $g_s > g_t$ determines the value of g, and hence the degree of undercoupling necessary to obtain the desired accuracy. For example, if $g_s=32$ and $g_t=10$, the error in Q_u is 2.7 percent.

III. EXPERIMENT

The general measuring procedure consists of the following. Samples of the incident and output signal are transmitted through two separate superheterodyne receiving systems. The two signals are introduced to a push-pull input of an oscilloscope and their signal levels adjusted for a null deflection. The ratio of the incident power to the transmitted power as a function of frequency is shown in Fig. 2. The frequency of the signal is adjusted so that P_t is a maximum and ω_0 may be determined. At this frequency, the input and output signals are balanced for a null deflection on the oscillo-

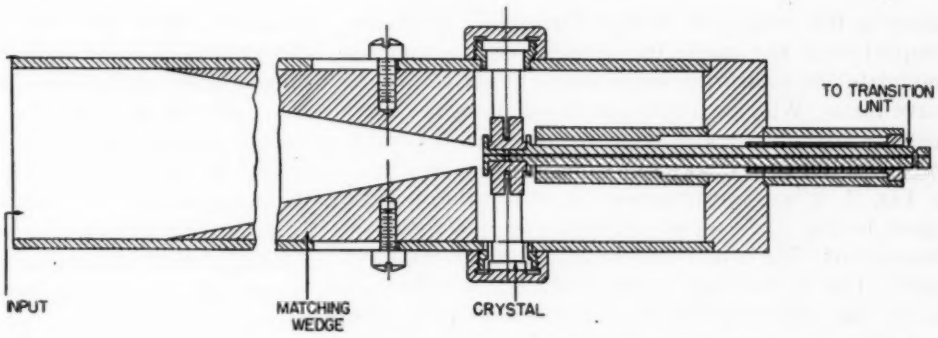


FIG. 4a. Wave-guide mixer.

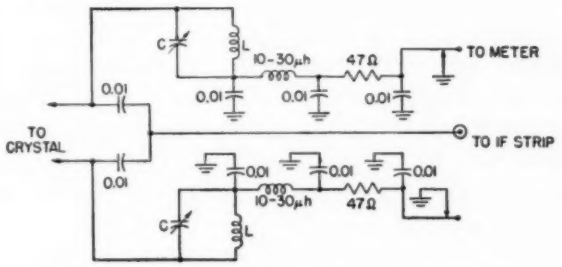


FIG. 4b. Transition unit from mixer to i.f. amplifier.

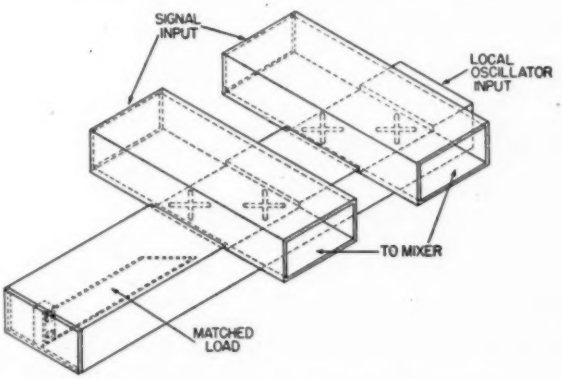


FIG. 4c. Double wave-guide directional coupler.

scope by adjusting the gain of the two receiving systems. At this point, by definition, the quantity γ is equal to one. When the attenuation in the receiving system of the transmitted power is decreased, the frequency must be changed to some value ω_1 in order to have a null signal. There exists another frequency ω_1' for which the same conditions hold true. The value of x is given by $\omega_1'-\omega_1$, and γ is obtained from the decrease in attenuation. If the attenuation is changed by 3 db, γ will have the value of two. By changing the value of the attenuation and measuring the proper frequency changes a complete set of data may be obtained. A block diagram of the equipment is shown in Fig. 3. The equipment consists of three sections, a microwave cavity with input and output couplings, a power-measuring section, and a frequency measuring section.

A. Power-Measuring Section

The microwave signals from the cavity are transmitted through a double 10-cm band wave-guide mixer unit in which the two signals have separate mixers but utilize the same local oscillator. The signals proceed through separate intermediate frequency (IF) amplifiers, video amplifiers, and cathode followers and are connected to a balanced input of an oscilloscope.

The double mixer is shown in Fig. 4a. It consists of two 10-cm band wave-guide mixers into which the local oscillator signal is inserted through a double wave-guide directional coupler. The local oscillator is frequency modulated by the oscilloscope sweep, which may or may not be externally triggered. Thus the method may be used for measuring the properties of a discharge in either steady state or transient condition. Each wave-guide mixer contains two crystals which are matched to the guide by wedges as shown. The two crystals serve a twofold purpose of increasing the mixer sensitivity and enabling the mixer to be matched. The IF signals from the crystals are added together through a transition unit from the mixer to the IF amplifier. The transition unit is shown in Fig. 4b and is designed so that the LC network resonates at the IF amplifier frequency, which for this case is 40 megacycles. The wave-guide directional coupler has been designed to have a coupling of 15 db and a directivity of 25 db and is shown in Fig. 4c.

Standard IF amplifiers may be used and the band width necessary will depend upon whether steady-state or transient measurements are to be performed. The important criterion is that the amplifiers be adjusted so that their band widths and IF frequencies are essentially identical. This is necessary in order to balance the signals to zero over the band width of the receiver. A precision wave-guide attenuator, calibrated to a precision of ± 0.05 db, is inserted in the section of guide containing the transmitted power and is used to produce the desired change in γ.

B. Frequency Measuring Section

The frequency is measured by means of a standard cavity wave meter using a sensitive receiver on the output to denote the wave-meter resonance. Rather than measuring directly the signal frequencies, it is more accurate and convenient to measure the difference between the signal frequency and the resonant frequency. A sample of the incident signal is transmitted through a magic tee into the wave meter. The side arms of the magic tee contain matched crystals which are connected to a General Radio signal generator, oper-

ating in the megacycle region. The signal generator, coupled with the magic tee, is utilized to amplitude modulate the measuring signal before it arrives at the wave meter. With no amplitude modulation, the wave meter is adjusted for the frequency at which the transmitted power is a maximum, which corresponds to ω_0 in Fig. 2. When the frequency is at a value ω_1, the signal to the wave meter is amplitude modulated so that the first side band is transmitted through the wave meter. The modulating frequency gives the difference $\omega_0-\omega_1$. In a like manner, $\omega_1'-\omega_0$ is obtained, and their sum gives the appropriate value of x. If the wave meter is adjusted exactly to the cavity resonance, the two readings will be identical. If they are not similar, their difference divided by two yields the deviation of the wave-meter reading from the true cavity resonance. Thus an accurate determination of the resonant frequency is possible within ± 10 kc.

IV. CONCLUSION

The advantages of this method are the following. The usefulness of plotting the data as a straight line is obvious. When the discharge admittance is measured, the method is most sensitive if the cavity is very much undercoupled, and hence the perturbation due to the coupling will be small. To the contrary, the resonant curve method has its greatest sensitivity when the cavity is very much overcoupled, and it is possible that the coupling perturbation may become important. In order to eliminate the possibility of disturbing the discharge with the measuring signal for the case when the measuring signal is not used to maintain the discharge, it is necessary to operate with a low-level signal. This can be done conveniently since a null method is employed and a lower signal-to-noise ratio may be tolerated than with the standing wave measurement technique. The problem of the effect of the slotted section and its probe on the measurements is no longer encountered. By the proper adjustment and calibration of equipment, an accuracy of ± 2 percent for the discharge admittance is feasible.

The authors wish to acknowledge the technical assistance of Mr. J. J. McCarthy.

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 8 AUGUST, 1953

Electrical Properties of Liquid Selenium I

H. W. HENKELS AND J. MACZUK
University of Pennsylvania, Moore School of Electrical Engineering, Philadelphia, Pennsylvania
(Received April 16, 1953)

The thermoelectric power of liquid selenium has been measured in the range 250–500°C. Values of activation energies calculated from the temperature dependences of thermoelectric powers and resistivities agree. This activation energy taken as the intrinsic activation energy (or the energy to break the Se—Se bond) is 2.31 ev calculated from data given by Henkels. The mobility of holes in the intrinsic region exceeds that of electrons and has a value $(8.4\times 10^6)/T^{\frac{1}{2}}$ cm²/volt sec, assuming electronic mass.

As a further check of the electronic nature of the conduction in the liquid and of theories of conduction in the hexagonal form, preliminary measurements of the thermoelectric power of pure selenium doped with small percentages of As, Sb, and Bi up to 4 percent were made. In the intrinsic region the specimens exhibited p type conduction, but at lower temperatures the sign of the thermoelectric powers reversed and values became quite large but negative. The crossover point depended qualitatively on the amount of impurity in the correct manner. The n type conduction is tentatively attributed to donor levels introduced by the normal trivalence of Group V_b elements. Two p electrons are considered bound in the selenium chain, the third ionized at high temperature.

INTRODUCTION

A REPORT of measurements of the conductivity of chemically pure liquid selenium was made by Henkels.[1] In the range of temperatures 200–500°C the resistivity was expressed by $\log_{10}\rho = A + (B/T)$. Average values for different selenium lots and melts of A and B were -3.81 and 5850°K. The maximum deviations from the averages were 10 percent and 3.4 percent, respectively. The resistivity was a function of temperature alone.

Various nonmetal impurities Cl_2, I_2, P lowered the resistivities and produced different values of B in different temperature ranges. Mercury addition in small amounts caused no change in either A or B, although that metal in the same quantities considerably influences the resistivity of solid hexagonal selenium.[2] It was assumed that the conduction was essentially electronic.

The present report describes measurements of the thermoelectric power and resistivity of pure selenium and preliminary studies of the influence of metals of Group V_b on the properties.

[1] H. W. Henkels, J. Appl. Phys. 21, 725–731 (1950).

[2] H. W. Henkels and J. Maczuk, "Electrical properties of selenium III, microcrystalline selenium metal doped" (to be published).

EXISTING DATA

Borelius *et al.*[3] conducted an extensive survey of liquid selenium. The thermoelectric powers and resistivities of selenium from a number of sources were measured. Their results show a rough agreement between values of activation energy (see later definition of the term) obtained from resistivity and thermoelectric power measurements. However, a large spread of values of this activation energy was obtained in different melts and in different strata of the same melt. It is assumed that unknown impurities were present in the selenium used in that investigation because highly reproducible values of liquid resistivity have been obtained in the work of Lizell[4] and in our own work.[1] (A slightly lower value of B was preferred by the former author.) Reference 2 demonstrates directly that, at least in the case of hexagonal selenium, the analysis of selenium should be extremely good (parts per million impurities) to avoid the influence of a number of metal contaminates. (The analyses for nonmetals O_2, Cl_2, Br_2, I_2 should be very low.)

TABLE I. Values of B and of $(\varphi+\zeta)$ from resistivity and thermoelectric power measurements.

Sample	B (°K)	$(\varphi+\zeta)_B$ (ev)	$D=(\varphi+\zeta)_D$ (ev)
1	5400	1.07	1.18
2	5800–5630	1.15–1.12	1.07
3	5200–5130	1.04–1.02	0.92

NEW DATA

Source of Selenium

Selenium cp of the Canadian Coppers Refiners, Ltd., was used in all experiments. This material was discussed in references 1 and 2.

Experimental Arrangements and Procedure

The cell used to contain the selenium during the measurements is shown in Fig. 1, together with details of the thermocouple well and the circuit arrangement. Platinum electrodes were used since these did not influence the properties of pure selenium in reference 1. The cell constant for resistivity determinations was obtained by calibration against a 0.01 N KCl solution. The entire arrangement was embedded in a sand bath centered in an electric oven. The selenium was heated for a few hours in a porcelain beaker at about 500°C, cooled, and poured into the preheated cell. The temperature of one arm of the cell was held 30°C above that of the other. After temperatures were steady, the resistances, thermoelectric powers, and temperatures were noted. The one-thousand cycle resistances were balanced in an ordinary bridge circuit. The thermovoltage produced was measured with a potentiometer using an indicator of sensitivity 6×10^{-10} amp/mm. Temperatures were measured with Chromel Alumel thermocouples in wells.

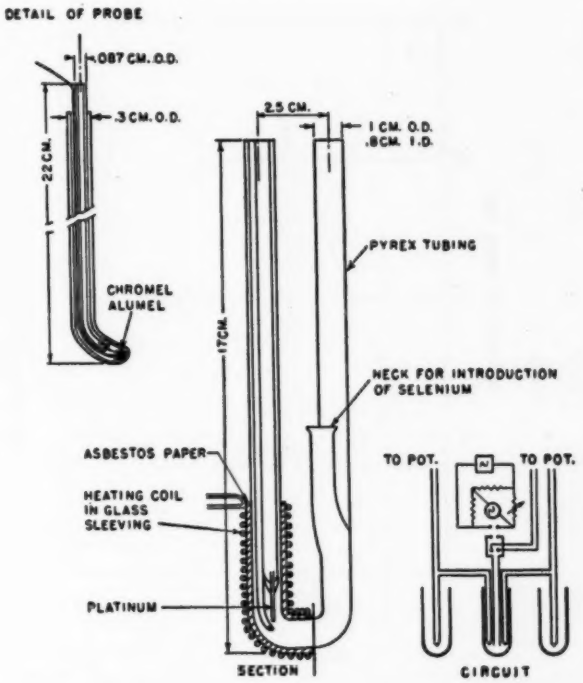


FIG. 1. Cell for measuring thermoelectric power in liquid selenium.

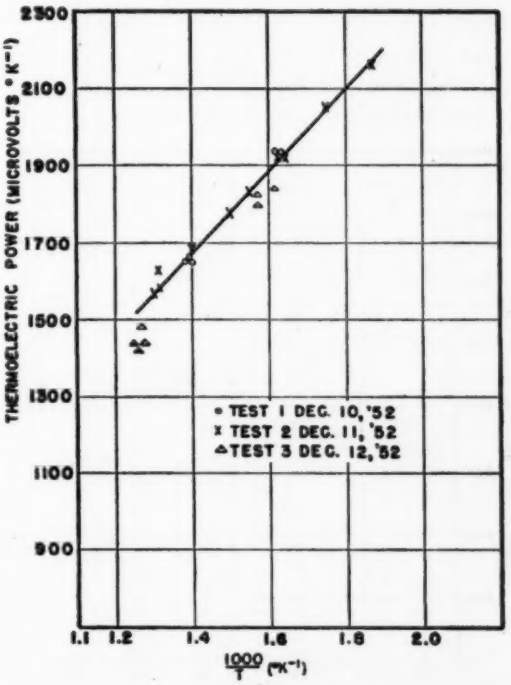


FIG. 2. Thermoelectric power of pure liquid selenium, ARQ brand—CCR, Ltd., heated 450°C—3 hours, tested in U-tube of design A, sample 2.

[3] Borelius, Pihlstrand, Andersson, and Gullberg, Arkiv. Mat. Astron. Fysik **30A**, No. 14 (1944).

[4] B. Lizell, Technical Report No. 47, Contract N5ori-7801, Laboratory for Insulations Research, Massachusetts Institute of Technology, September (1951).

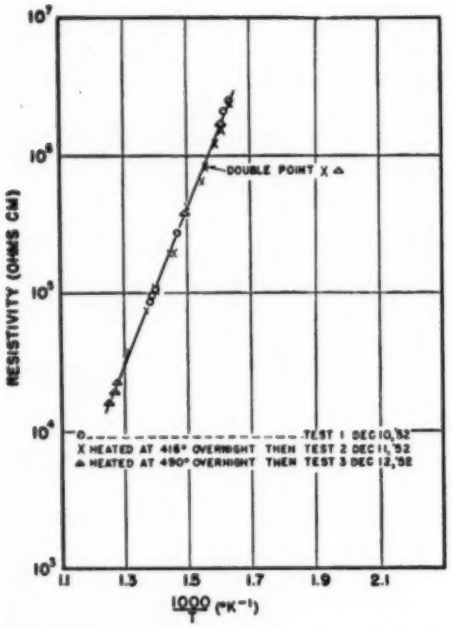

RESISTIVITY (OHMS CM)
DOUBLE POINT X △
X HEATED AT 416° OVERNIGHT THEN TEST 2 DEC 11,'52
△ HEATED AT 490° OVERNIGHT THEN TEST 3 DEC 12,'52
1000/T (°K⁻¹)

FIG. 3. Resistivity of pure liquid selenium, ARQ brand—CCR, Ltd., heated 450°C—3 hours, tested in U-tube of design A, sample 2.

Data—Pure Selenium

Figures 2 and 3 gives representative data on the thermoelectric powers and resistivities of a sample tested continuously for three days. The first night of this period the selenium was held at 416°C, the second night at 490°C. Various data are presented in Table I in terms corresponding to empirical and theoretical expressions given below, the symbols having the meanings:

ρ = resistivity (ohm-cm), a, A, B, C, D = constants,
T = absolute temperature (°K),
q = electronic charge (coulombs),

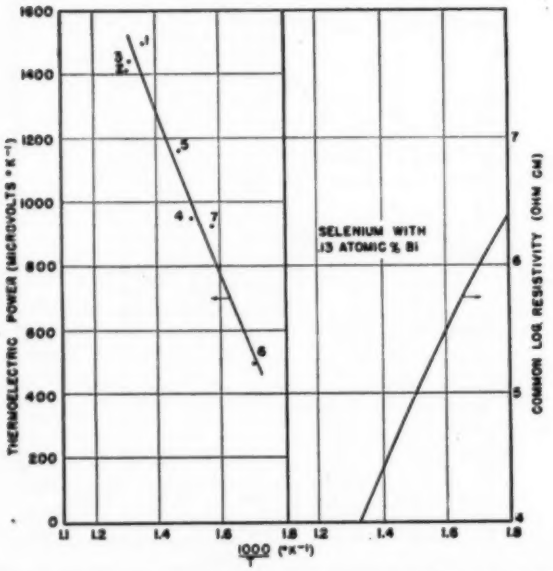


FIG. 4. Resistivity and thermoelectric power of liquid selenium with impurity of bismuth, ARQ brand—CCR, Ltd.

n_1, n_2 = number densities of electrons and holes, respectively (cm^{-3}),
b_1, b_2 = mobilities of electrons and holes, respectively (cm^2/volt sec),
Q = thermoelectric power per degree (volts/°K),
k = Boltzmann constant (joules/°K),
φ = width of forbidden zone (a positive number in electron volts),
ζ = Fermi level measured from an arbitrary zero at the bottom of the conduction band (positive values increasing with energy in electron volts),
h = Plank's constant (joule sec),
m_1, m_2 = masses of electrons and holes, respectively (kg).

It is found that $\log_{10}\rho$ can be expressed empirically as

$$\log_{10}\rho = A + (B/T), \tag{1}$$

the thermoelectric power as

$$Q = C + (D/T). \tag{2}$$

On the basis of a simple intrinsic model in the Wilson terminology of band structure the resistivities,

$$1/\rho = q(n_1 b_1 + n_2 b_2), \tag{3}$$

TABLE II. Values of A and C from resistivity and thermoelectric power measurements.

Sample	A	C
1	−3.33	-0.066×10^{-3}
2	−2.95	$+0.165\times10^{-3}$
3	−2.18	0.293×10^{-3}

and the thermoelectric powers,

$$Q = -\frac{k}{q}\left\{\frac{2(n_2b_2 - n_1b_1)}{(n_2b_2 + n_1b_1)} + \frac{n_1b_1}{n_2b_2 + n_1b_1}\left(\ln\frac{n_1h^3 10^6}{2(2m_1\pi kT)^{\frac{3}{2}}}\right) - \frac{n_2b_2}{(n_2b_2 + n_1b_1)}\left(\ln\frac{n_2h^3 10^6}{2(2m_2\pi kT)^{\frac{3}{2}}}\right)\right\}, \tag{4}$$

(these written with the assumptions of the existence of scalar mean free paths and masses, both independent of energy) can be simplified to the forms,

$$\log_{10}\rho = \log_{10}\frac{h^3 10^6}{qa(2)(2m_2\pi k)^{\frac{3}{2}}} + \frac{q(\varphi+\zeta)}{kT}\log_{10}e, \tag{5}$$

$$Q = \frac{2k}{q} + \frac{\varphi+\zeta}{T}, \tag{6}$$

with the additional assumptions that the mobility of holes considerably exceeds that of electrons and varies as $b = a/T^{\frac{3}{2}}$.

Hence,

$$A = \log_{10}\frac{h^3 10^6}{qa(2)(2m_2\pi k)^{\frac{3}{2}}} \tag{7}$$

$$B = \frac{q(\varphi+\zeta)}{k}\log_{10}e \tag{8}$$

$$C=2k/q \tag{9}$$

$$D=\varphi+\zeta. \tag{10}$$

Values of B, $(\varphi+\zeta)_B$, $(\varphi+\zeta)_D$ taken from the slopes of the $\log_{10}\rho$ and Q *versus* $1/T$ curves are listed in Table I.

Sample 3 was measured in a cell which had been previously used to test antimony-doped specimens so that the distinctly lower value is attributed to the presence of traces of antimony remaining after careful cleaning. The differences in values of $(\varphi+\zeta)_B$ are to be attributed to experimental errors in noting thermoelectric power and to the presence of impurities (including oxygen) in the specimens. The precision in the determination of thermoelectric power was not sufficiently high to justify recalculations of the values of $(\varphi+\zeta)$ on the basis of different ratios b_2/b_1. Values of A and C are given in Table II.

The C value averages around $2k/q=0.17\times10^{-3}$ ev required by the theory. From A of sample 2 a value of mobility can be obtained assuming electronic mass for holes. However, the experimental difficulties in measuring the thermoelectric power and resistivity simultaneously and the averaging of resistances over temperature gradients not exactly known makes the previous precise measurements of resistivity, reference 1, more reliable in calculating properties of the pure material. The good agreement of values of activation energy obtained from thermoelectric power and resistivity measurements then encourages the calculation of the intrinsic properties of pure selenium from values of $A=-3.81$, $B=5850°$K. A later section of this report will present additional data to support the electronic interpretation of the resistivity.

From (7) b is calculated to be

$$b=\frac{8.37\times10^{6}}{T^{\frac{3}{2}}}\ \text{cm}^2/\text{volt sec.}$$

From (8) the width of the forbidden band in liquid selenium is

$$\varphi=2.31\ \text{ev.}$$

Data—Selenium Doped with As, Sb, Bi

It has been concluded from a study of the influence of metals on hexagonal selenium that materials electropositive with respect to selenium enhance electrostatic barriers existing in the purest selenium by producing donor levels. Resistivities and thermoelectric powers both increase indicating the depletion of surface layers of the crystallites. However, with impurities of metals up to 0.1 percent, the thermoelectric power remains positive and large. The fact that the thermoelectric power does not reverse sign has been attributed to the general insolubility of impurities in selenium. The metals of Group V_b are particularly active in altering the electrical properties of the solid material. It was reasoned that in liquid and undercooled selenium the solubility problems would be avoided. Therefore the resistivities and thermoelectric powers of liquid selenium doped with bismuth, antimony, and arsenic were measured in a preliminary study.

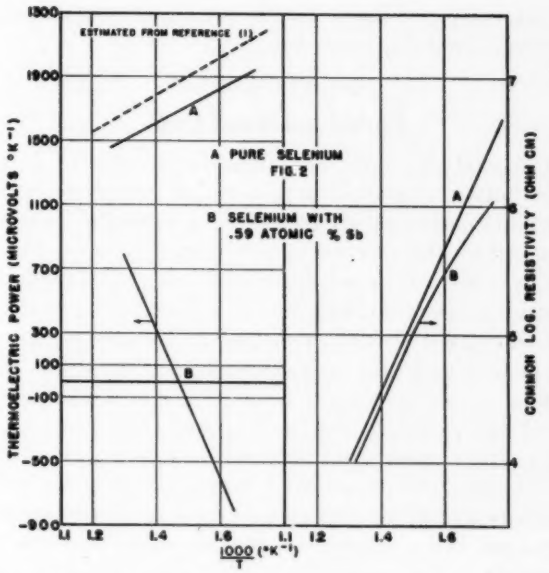

ESTIMATED FROM REFERENCE (1)
A PURE SELENIUM FIG. 2
B SELENIUM WITH .59 ATOMIC % Sb
THERMOELECTRIC POWER (MICROVOLTS °K⁻¹)
COMMON LOG. RESISTIVITY (OHM CM)
1000/T (°K⁻¹)

FIG. 5. Resistivity and thermoelectric power of liquid selenium with impurity of antimony, ARQ brand—CCR, Ltd.

Figure 4 presents data for a specimen doped with 0.13 atomic percent Bi, Fig. 5, that for one with 0.59 atomic percent Sb, Fig. 6, for a collection of As-doped samples. It is noted that all three of these materials make the sign of the thermoelectric power reverse. Unfortunately experimental difficulties prevented the extension of the studies to lower temperatures. In the cases of specimens doped with antimony or bismuth, the data were fairly reproducible on repeated measurement. The results are of a preliminary nature and more reliable measurements will be made in extended tem-

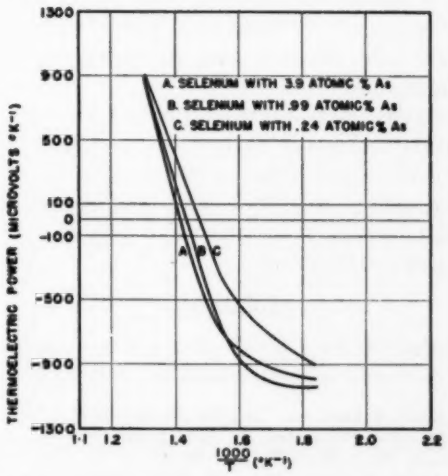

A. SELENIUM WITH 3.9 ATOMIC % As
B. SELENIUM WITH .99 ATOMIC% As
C. SELENIUM WITH .24 ATOMIC% As
THERMOELECTRIC POWER (MICROVOLTS °K⁻¹)
1000/T (°K⁻¹)

FIG. 6. Thermoelectric power of liquid selenium with impurities of arsenic, ARQ brand—CCR, Ltd.

perature regions before donor activation energies and electronic mobilities are calculated.

CONCLUSIONS

Forbidden Band Gap

The good agreement of activation energies calculated from both thermoelectric power and resistivity measurements indicates an electronic conduction process. Support to the electronic interpretation is given by the data on *n* type selenium produced by doping the chemically pure material with metals of Group V_b. The calculated activation energy can be used to extend the resistivity values to room temperature. When this is done rough agreement is obtained with the orders found by Keck[5] in the measurement of the charge decay from amorphous selenium coatings. His data do not show the exponential dependence of resistivity on temperature, but the range was very small and the technique not adapted to precise measurements of resistivity.

Reports of studies of the photoconductivity and optical absorption in liquid selenium are not known to us. The resistivity of chemically pure selenium at 150°C (i.e., undercooled liquid) depends on temperature in the same manner as does the liquid. (Precise measurements are difficult at lower temperatures because of the extremely high resistivities and the large effects of any traces of impurities.) Data on the optical properties of undercooled selenium at room temperature are available. Weimer[6] finds a spectral response peaked in the blue-green portion of the spectrum with very low response in the red compared with the peak sensitivity to red light in hexagonal selenium. The blue sensitive photoconductivity was suggested as being largely of a primary nature. Hole conduction was found to be predominant over electron conduction by that author and by Pensak.[7] Gilleo[8] found that the absorption edge in amorphous selenium at 295°K corresponded roughly to the Se—Se bond strength 2.5 ev given by Pauling.[9] Photoconductivity at the same temperature began at a wavelength of 5600A about equal to that given by Weimer compared with about 7600A for the hexagonal form.

With the electronic interpretation of the conductivity, the thermal ionization energy required to break the Se—Se bond in the liquid or alternately the intrinsic activation energy is 2.31 ev, and this agrees with the optical data within the rather poor precision of interpretation of the latter.

Mobilities

The value of $8.4\times10^6/T^{\frac{1}{2}}$ cm²/volt sec for mobility of holes in liquid selenium between 500°C and 250°C is very high compared with those measured in the hexagonal forms (the highest value found at 200 mc was about 30 cm²/volt sec). It will be recalled that electronic mass was assumed in all calculations so that the mobility should be written as

$$b=\frac{b_{m=m_e}}{(m_2/m_e)^{\frac{1}{2}}}$$

From data in Fig. 5 it is estimated that the mobility of electrons is extremely small compared with this value.

Donor Levels

We have postulated that arsenic, antimony, and bismuth alter the properties of hexagonal selenium markedly by the formation of localized donor levels enhancing electrostatic barriers. In liquid selenium the elements in question should be soluble in the amounts used so that a distribution of such levels throughout the selenium was expected. With the greater solubility, the thermoelectric power does in fact reverse below the intrinsic region. As yet the results are of a preliminary nature and not extensive enough to warrant calculation of activation energies for donor levels or mobilities of electrons.

The metals of Group V_b increase the viscosity of selenium at low temperatures presumably by the formation of additional cross links between selenium chains by means of the extra *p* electron. The action of these metals in forming donor levels is again tentatively attributed to their trivalence. Whereas in tetrahedral structures the five electrons would be expected to form hybrid bonds, in the chain structure of selenium, two of the *p* electrons bond in the chain, the third may be ionized at higher temperatures.

In all the above, calculations have been based on models arising in the band theory of solids. The terms used in the later and the quantitative theory are applicable to perfectly periodic structures. In considering the electrical properties the degree of order required of a solid for an exact correspondence between theory and experiment might appear to be in the range of the mean free path of the current carriers. The free path is quite small in most metals and semiconductors. On the other hand, the description of electrical properties in molecular terms must give essentially the same results with important modifications only in terminology and in quantitative detail. The activation energy in the electronic process is equally well described as the energy corresponding to the forbidden band gap or the energy to break a bond in the selenium homopolar lattice. Calculations of the carrier densities necessary in estimating mobilities involve constants which may be somewhat in error when applied to a disordered lattice. However, the mobility results would seem to be correct to within small factors provided appropriate masses of carriers are used.

[5] P. H. Keck, J. Opt. Soc. Am. **41**, 53–55 (1951).
[6] P. K. Weimer, Phys. Rev. **79**, 171 (1950).
[7] L. Pensak, Phys. Rev. **79**, 171 (1950).
[8] M. A. Gilleo, J. Chem. Phys. **19**, 1291 (1951).
[9] L. C. Pauling, *The Nature of the Chemical Bond* (Cornell University Press, Ithaca, 1940).

Letters to the Editor

Schlieren Photographs of Sound Fields

DANIEL S. SCHWARTZ AND ANTHONY L. RUSSO
Cornell Aeronautical Laboratory, Inc., Buffalo 21, New York
(Received March 4, 1953)

IN recent years considerable interest has been aroused in the subject of aerodynamically generated sound because of the intense noises produced by high velocity jet streams.[1,2] While a great deal of attention has been devoted to the measurement of the intensity of these fields, it has been only recently that theoretical investigations of the mechanism of sound generation have been initiated.[3] Photographs of sound fields taken at Cornell Aeronautical Laboratory with a sensitive schlieren system may shed more light on the sound-generating mechanisms of turbulent jets.

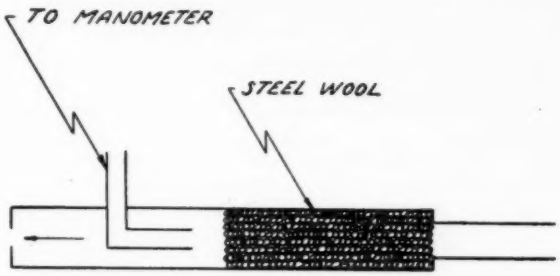


FIG. 1. Low Mach number jet geometry.

Schlieren photographs were taken of the sound field of both supersonic and subsonic jets in a range of Mach numbers from 0.7 to above 1. For supersonic testing, a simple $\frac{1}{4}$ in. o.d. copper tube was attached to a nitrogen bottle from which the gas issued at a gauge pressure of approximately 200 psi, producing a choked jet. The subsonic jet as shown in Fig. 1, was produced by a 1 in. brass tube with a $\frac{1}{4}$ in. orifice at the exit end. In order to prevent line pulsations which may have modified the free jet phenomena, a damping chamber consisting of steel wool packing was inserted into the 100 psi air supply. To determine the total pressure of the jet, a pitot tube was inserted immediately downstream from the settling section.

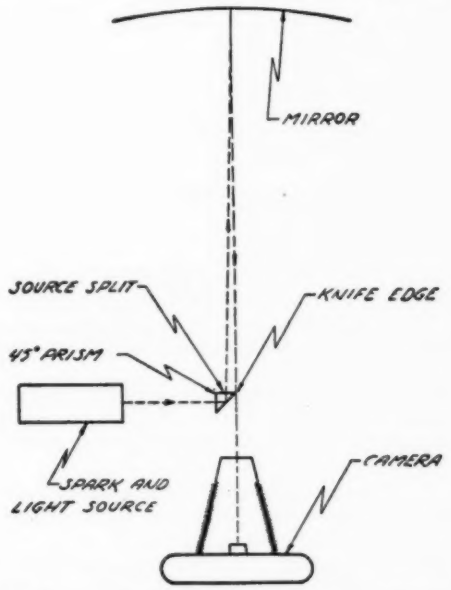


FIG. 2. Schematic diagram of the apparatus used to investigate sound fields.

The optical system used (Fig. 2), suggested by A. Hertzberg, was of the single-mirror, double-pass type, utilizing a small right-angle prism to reduce the parallax usually encountered in the conventional double-pass schlieren system. One corner of the prism, used as the knife edge, was defined by black masking tape on one of its faces. The system was equipped with a 10v-7.5 amp focusing lamp and a spark source of approximately one microsecond duration. A 6-foot focal length, 12-inch diameter spherical mirror reflected the light to the knife edge, and into the camera through a single 0.75-meter focal length lens. The system was aligned with the filament, and preliminary photographs were taken with a standard polaroid camera to check the alignment. Upon receiving satisfactory focus and sensitivity, photographs were then taken

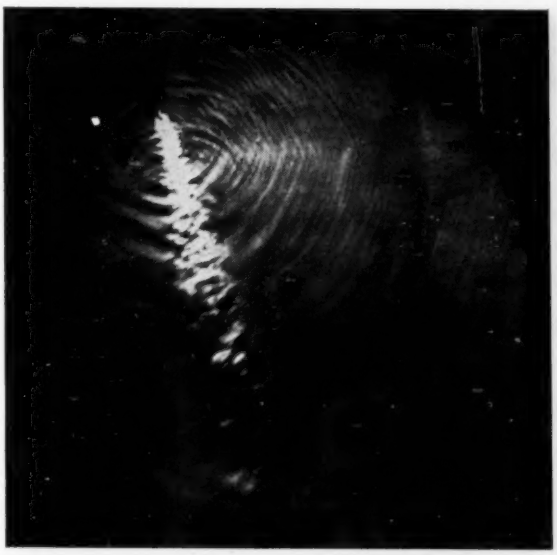


FIG. 3. Schlieren photograph at high sensitivity of the sound field associated with a supersonic nitrogen jet.

FIG. 4. Schlieren photograph at moderate sensitivity of the sound field associated with a supersonic nitrogen jet.

FIG. 5. Schlieren photograph of the sound field associated with a subsonic air jet.

with several types of commercial films of which Kodak contrast process ortho gave the best results.

Figures 3 and 4 are schlieren photographs of the supersonic nitrogen jet taken with different sensitivities. The mixing zone in this jet was always found to be turbulent. From the photographs there appear to be at least two sources of sound emission, one at the jet exit and one in the turbulent zone.

Figure 5 shows the sound field of the damped air jet issuing at a velocity of Mach 0.7. It is to be expected that the noise field from the turbulent zone of this subsonic jet would be less intense than that of the supersonic jet, since turbulence at low Mach numbers is a poor generator of sound.

This investigation will be extended to jets of larger diameters. It is hoped that this study may lead to an experimental method of verifying Lighthill's theory of aerodynamically generated sound and to practical methods of reducing the intensity of the sound produced by jet streams.

[1] H. R. Saffell and H. S. Hansen, USCAL Report 6-1-2 (August 5, 1948).
[2] L. W. Lassiter, National Advisory Committee on Aeronautics, NACA TN 2756 (August 1952).
[3] J. J. Lighthill, Proc. Roy. Soc. (London) **A211**, 1107, 564–587 (1952).

Effect of Cold Work and Anneal on Resistivity of Alpha Brass

HERBERT I. FUSFELD
Frankford Arsenal, Philadelphia, Pennsylvania
(Received April 8, 1953)

DURING a study of cold work and subsequent anneals in polycrystalline 70-30 brass, unusual resistivity effects were observed. These are now under further study. The writer's attention has been called to results of Masima and Sachs[1] on changes in density of monocrystalline 70-30 brass, repeated in this laboratory by D. J. Murphy on polycrystalline brass, similar to the observed resistivity changes. A comparison of these results is given in this letter.

Alpha-brass rods of a high-grade commercial variety, grain size less than 0.020 mm, were pulled in tension in three groups: 1 percent, 10 percent, and 20 percent elongation. Ends were cut off to leave 15-in. specimens in uniformly stained condition. The five specimens within each group were annealed at separate temperatures from room temperature to about 245°C, below the recrystallization point.

The property originally under study was internal friction, and each rod was kept at the annealing temperature for approximately one day to allow recovery of the original internal friction value prior to cold work. Later, resistivity measurements were made at room temperature, with time between cold work and measurement as follows:

Anneal temperature	28°C	93°C	149°C	205°C	246°C
Time in months	8	7	4	3	¾

Plotted in Fig. 1 are curves of residual resistivity *vs* temperature of anneal or recovery for specimens given 1 percent, 10 percent, and 20 percent strain, respectively. Also shown is the data for density *vs* temperature of anneal obtained by Masima and Sachs for single crystals of similar material. However, their specimens were annealed at the various temperatures for one half-hour.

The resistivity results show these characteristics:

(1) Specimens recovering at room temperature reached the same value.

(2) The curves lie above each other in the same order as the strain.

(3) A maximum and minimum is found in each curve at about 100°C and 180°C, respectively, compared to similar phenomena in the density data at about 200°C and 400°C.

Concentrating on this third feature, assume that the underlying mechanism for both density and resistivity effects is a rate process proportional to the quantity $te^{-Q/RT}$, where

t = time of anneal at temperature T, T = temperature of anneal in °K, R = gas constant, and Q = activation energy.

If one considers that both effects have the same average activation energy, then the time and temperature are related by the

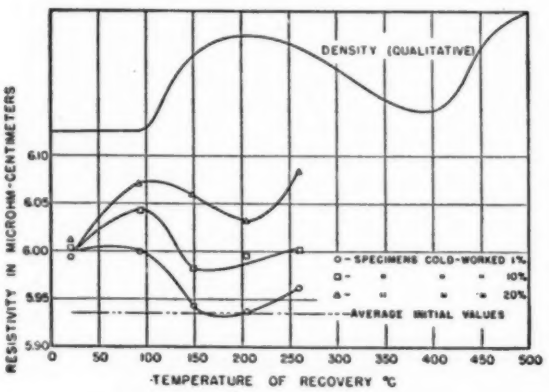


FIG. 1. Changes in resistivity and density of cold-worked 70-30 brass as functions of annealing temperature. All measurements made at room temperature.

expression: $te^{-Q/RT}$ = constant. The ratio of t (resistivity)/t (density) is about 50. Hence, using the quantities given in (3) above, the shift in maxima would be accomplished by a value of Q = 21 000 calories/mole, while the shift in minima would occur for Q = 17 000 calories/mole.

Thus the assumption of approximately 19 000 calories/mole for both processes would allow the density data to transform to the same scale as the resistivity data, so that both maxima and minima coincide.

The concept deduced from these results is that several mechanisms operate simultaneously during the recovery process, which compete in their effects in raising or lowering the resistivity and density. Among these are (a) Anchoring of dislocations by impurities,[2] lowering resistivity; (b) mutual annihilation of dislocations, greatly lowering resistivity; (c) migration of dislocations from anchoring impurities, particularly at higher temper-

atures, raising resistivity; and (d) dispersion of vacancy clusters, raising resistivity.

The relative quantitative effects are not well established. The pertinent activation energies, dislocation densities, and crystal structure for a given metal would influence the position of such maxima and minima and determine whether they occur at all.

[1] F. Seitz, Advances in Physics I, No. 1, 43–90 (1952).
[2] A. H. Cottrell, *Report of Conference on Strength of Solids* (Physics Society, London, 1948), p. 30.

A Simpler Explanation for the Observed Shot Effect in Germanium Filaments

A. VAN DER ZIEL
Electrical Engineering Department, University of Minnesota, Minneapolis, Minnesota
(Received April 16, 1953)

TRUE shot effect in n-type germanium filaments was first observed by Herzog and van der Ziel.[1] Mattson and van der Ziel[2] extended the measurements to higher frequencies and gave a possible explanation of the effect. It was found that a fluctuation in the number of holes could not explain the magnitude of the observed noise. A fluctuation in the number of electrons could explain the observed shot noise if it was assumed that the electrons came from traps at a depth of 0.23 ev. This explanation is in itself not impossible, since traps at that depth have been found by other observers, but the observed Hall effect of the filaments seems to exclude the possibility that most electrons come from traps.[3]

The following explanation does not violate the Hall effect data and is simpler. Consider a slab of semiconducting material between plane electrodes at distance d and let a voltage V_0 be applied. Let N_e be the number of free electrons in the sample and N_h the number of free holes. Let it be assumed that hole traps and electrons traps play only an unimportant part—that is, most electrons come from donor levels. Since nearly all these donor levels are ionized at room temperature, they cannot be a source of fluctuations. Some of the electrons and all the holes come from hole-electron pair creation. The processes of hole-electron pair creation and hole-electron recombination produce a fluctuation in the number of holes and an equal fluctuation in the number of electrons, so that

$$(N_e-\bar{N}_e)=(N_h-\bar{N}_h);\quad \overline{(N_e-\bar{N}_e)^2}=\overline{(N_h-\bar{N}_h)^2}=\bar{N}_h. \tag{1}$$

Since the fluctuating number of holes produces an equal fluctuating number of electrons, the total fluctuating current will be much larger than when only the number of holes fluctuated. The fluctuating current in the external lead is,

$$I(t)=(eV_0/d^2)(\mu_e+\mu_h)(N_h-\bar{N}_h), \tag{2}$$

where e is the electron charge and μ_e and μ_h the electron and hole mobilities, respectively. The conductance g of the sample is

$$g=(e/d^2)(\mu_e\bar{N}_e+\mu_h\bar{N}_h), \tag{3}$$

and the intensity of the thermal noise is $4kTg$.

The shot noise intensity is, therefore,

$$w(f)=4\int_0^\infty \overline{I(t)I(t+s)}\cos\omega s\,ds=4\bar{I}^2\tau/(1+\omega^2\tau^2), \tag{4}$$

since

$$\overline{(N_h-\bar{N}_h)_t(N_h-\bar{N}_h)_{t+s}}=\bar{N}_he^{-s/\tau}, \tag{4a}$$

where τ is the lifetime of the holes. The shot noise ratio n_s (shot noise intensity over thermal noise intensity) is, therefore,

$$n_s=\frac{e}{kT}\frac{V_0^2}{d^2}\frac{(\mu_e+\mu_h)^2}{\mu_h}\frac{\sigma_h}{\sigma_e+\sigma_h}\frac{\tau}{1+\omega^2\tau^2}, \tag{5}$$

where σ_e is the electron conductivity and σ_h the hole conductivity of the sample. The shot noise ratio in Herzog and van der Ziel's sample at low frequencies ($\omega\tau\ll1$) was 54 for $V=30$ volts and $d=0.52$ cm, and the life time deduced from the noise spectrum was 1.1×10^{-6} sec. On substituting this in Eq. (5) we get $\sigma_h/(\sigma_e+\sigma_h)=0.023$. Combining this result with the measured resistivity of 16 ohm-cm of the sample, it is possible to calculate the resistivity of intrinsic germanium (compare Eq. (6)). The resulting value is 50 ohm-cm, in excellent agreement with the accepted value of 47 ohm-cm at 300°K.[4]

An indirect verification of the new shot noise theory is supplied by the fact that a sample having a resistivity of 1.6 ohm-cm did not show any measurable shot effect at the same dc power level. For if σ_{ie} and σ_{ih} are the intrinsic electron and hole conductivities of the material, then

$$\sigma_e\sigma_h=\sigma_{ie}\sigma_{ih}\quad\text{or}\quad \sigma_h/(\sigma_e+\sigma_h)=\sigma_{ie}\sigma_{ih}/(\sigma_e^2+\sigma_{ie}\sigma_{ih}), \tag{6}$$

so that $\sigma_h/(\sigma_e+\sigma_h)$ is about 100 times as small as for the other sample. Since V_0^2 is 10 times as small for the same power level, the shot noise ratio should be about 1000 times as small as for the other sample. This should indeed be unobservable.

The foregoing mechanism thus can explain the observed shot noise. Montgomery[5] has shown that the *excess* noise of the filaments (that is the noise component with a $1/f$ spectrum) is also caused by the holes, so that both effects seem to be the result of a similar cause.

This type of shot effect should only occur in samples that are nearly intrinsic; this is perhaps the main reason why it has not been observed before. The noise should increase rapidly with increasing temperature. A measurement of the noise ratio as a function of temperature will test whether the foregoing explanation is correct.

The author is indebted to H. C. Montgomery of Bell Telephone Laboratories, whose helpful criticism stimulated the search for a better theory of the observed shot effect.

[1] G. B. Herzog and A. van der Ziel, Phys. Rev. **84**, 1249 (1951).
[2] R. H. Mattson and A. van der Ziel, J. Appl. Phys. **22**, 222 (1953).
[3] H. C. Montgomery (private communication).
[4] E. M. Conwell, Proc. Inst. Radio Engrs. **40**, 1327 (1952).
[5] H. C. Montgomery, Bell System Tech. J. **31**, 950 (1952).

Theory of the Flicker Effect

T. B. TOMLINSON AND W. L. PRICE
University of Southampton, Southampton, England
(Received April 13, 1953)

THERE has been considerable interest recently in the mechanism underlying low-frequency fluctuations of emission current from oxide-coated cathodes and in the similar effect of contact noise in semiconductors. In the case of the oxide-coated cathode, opinion seems to be divided between a surface and a volume effect, and it is with the former that we are concerned here. The first theory of this kind is that of Schottky,[1] who regarded the phenomenon as a fluctuation in the number of emission centres with smaller outer work functions (flicker effect). In an extension of this theory Macfarlane[2] assumed that the surface changes were due to the formation of a dipolar layer, which lowered the potential barrier in localized areas. Other modifications of Schottky's theory have appeared elsewhere.[3,4]

Both the above theories would appear to apply to the diode working under conditions of emission saturation; it is interesting to reconsider them for different operating conditions. In a previous paper,[5] investigations into the low-frequency fluctuations of a diode operated under retarding field conditions were described; the results show that generally the mean-square current fluctuations increase as the square of the mean dc current, until space-charge conditions set in, although lower-power indices are found. As space charge appears the increase of current fluctuations continues at a reduced rate, the transition being smooth and continuous.

Under true retarding field conditions there is no potential minimum due to space charge in the inter-electrode region; the point of lowest potential is at the anode, and the anode current is then dependent only on the initial velocities of the emitted electrons, the anode-cathode potential difference, and the work function of the anode material. The current cannot be changed by alteration of the potential barrier at the emitting surface unless such alteration causes the height of the barrier to rise above that at the anode surface. A reduction in the height of the cathode surface barrier will have no effect. Hence it is clear that neither of the above theoretical models would account for current fluctuations under retarding field conditions. Needless to say, statistical fluctuations have been neglected in this argument.

On the other hand, it is reasonable to assume that for some of the patches of the heterogeneous cathode surface the height of the barrier does approach the height of the potential hill at the anode. In this case an instantaneous increase of the surface barrier at a patch above the level of the anode potential barrier will lead to a reduction of the anode current. The inference is that only increases of cathode outer work function, for instance those due to the presence of negative ions on the surface, can be the cause of emission fluctuations under retarding field conditions.

As the negative anode-cathode potential difference is reduced, the operating conditions gradually become space-charge limited. This means that the point of minimum potential moves from the anode into the inter-electrode space, and that the height of the potential barrier (now associated with the space charge) is gradually being reduced as the anode is made more positive. Consequently, the mechanism of emission fluctuation is unchanged but the magnitude of the fluctuations will increase. When conditions of complete emission saturation are reached, decreases of the outer work function of the cathode surface may also have the effect of causing variations of emission. In other words, the two above-mentioned theories are restricted to saturation conditions only.

In a second theory of contact noise, which is specially applicable to the rmionic emission from a semiconductor, Macfarlane[6] introduces an increased potential barrier at the surface due to the presence of negative ions. The surface of the semiconducting crystal is assumed to be covered by a partial film of adsorbed atoms, which become negative ions by the capture of electrons from the crystal. The height of this barrier is proportional to the concentration of negative ions which are assumed to move at random over the surface by a diffusion process. The emission is assumed to take place at localized patches on the surface, and variations in the height of the potential barrier due to changing concentration of ions in a patch lead to fluctuations in the emission current.

Thus it is assumed that the current j through a patch depends on the number n of adsorbed ions within the patch as $\exp(an)$ where a is a negative constant. This implies that both increasing and decreasing values of n are effective in changing the emission current, i.e., the mean emission current j_0 must be saturated or temperature limited, and no space-charge effects are considered.

Application of the theory to a thermionic diode must take into account the fact that the emission centers or patches will have different work functions. Nevertheless, for a given patch the potential barrier at the oxide-vacuum junction is a combination of the outer work function and the additional barrier due to negative ions. Once an electron has passed the barrier it will automatically proceed to the anode so that both a rising and a falling barrier height will influence the anode current. The current through any patch is given by

$$j=j_0\exp[a(n-\bar{n})],$$

where n is the number of negative ions within the patch and is distributed normally over the cathode surface with mean value $\bar{n}$ and standard deviation $(\bar{n})^{\frac{1}{2}}$. Any intermediate barriers due to interface layers, etc. should have no effect at normal dc currents and are omitted. The following is a modification of this theory to take into the account the changed circumstances relevant to retarding field operation.

An instantaneous increase in the number of adsorbed negative ions will lead to a reduction of anode current, if the instantaneous height of the potential barrier at the surface of the patch exceeds the potential hill at the anode; a reduction of the surface barrier height will have no effect. The mean-square value $\langle\delta j\rangle_{Av}^2$ of the anode current fluctuations is assessed on this basis. Furthermore, due to the one-way action of these fluctuations the mean anode current $\bar{j}$ will be less than that value corresponding to the initial velocity distribution and the height of the potential hill at the anode. Therefore its value is calculated so that $\langle\delta j\rangle_{Av}^2$ may be expressed in terms of $\bar{j}$.

The component of anode current caused by emission from the patch is

$$j_0\exp[a(n-\bar{n})] \quad \text{for} \quad n>n_0$$

or

$$j_0\exp[a(n_0-\bar{n})] \quad \text{for} \quad n<n_0,$$

where $|an_0|$ is the height of the potential hill at the anode. Hence the mean anode current is

$$\bar{j}=\int_{n_0}^{\infty} j_0\exp[a(n-\bar{n})](2\pi\bar{n})^{-\frac{1}{2}}\exp\left[\frac{-(n-\bar{n})^2}{2\bar{n}}\right]dn$$
$$+\int_{-\infty}^{n_0} j_0\exp[a(n_0-\bar{n})](2\pi\bar{n})^{-\frac{1}{2}}\exp\left[\frac{-(n-\bar{n})^2}{2\bar{n}}\right]dn$$
$$=\tfrac{1}{2}j_0\left[1-\operatorname{erf}\left(\frac{-a\sqrt{\bar{n}}}{\sqrt{2}}+\frac{n_0-\bar{n}}{\sqrt{2}\sqrt{\bar{n}}}\right)\right]\exp\left(\frac{a^2\bar{n}}{2}\right)$$
$$+\tfrac{1}{2}j_0\left[1+\operatorname{erf}\left(\frac{n_0-\bar{n}}{\sqrt{2}\sqrt{\bar{n}}}\right)\right]\exp[a(n_0-\bar{n})].$$

We now define the variables,

$$x=-a\left(\frac{\bar{n}}{2}\right)^{\frac{1}{2}}, \quad y=(n_0-\bar{n})(2\bar{n})^{-\frac{1}{2}}.$$

For a given emission characteristic, x is fixed and y is a linear function of the retarding field. Then,

$$\bar{j}=\tfrac{1}{2}j_0\{[1-\operatorname{erf}(x+y)]\exp(x^2)+[1+\operatorname{erf}y]e^{-2xy}\}.$$

For $y>2$, $\bar{j}=j_0\exp(-2xy)$ to an accuracy better than $\frac{1}{2}$ percent for all values of x.

If $x\ll 1$, $\bar{j}\simeq j_0\exp(-2xy)$ for $y>0$, i.e., for retarding field conditions. The retarding field current is known to be governed by a work function of the order of an electron volt. The magnitude of xy is then of order 10; as the number of negative ions involved in the action must be high, we conclude that y is large, a is small, and x is small.

The fluctuation of anode current is given by

$$\delta j=j_0\exp[a(n-\bar{n})]-\bar{j} \quad \text{for} \quad n>n_0.$$

Following Macfarlane, one obtains the autocorrelation function $f(\tau)$ of the noise current in terms of the probability $P(\tau)$ that a negative ion in any patch will have emerged from it in time τ.

For $x\ll 1$ the approximate result is

$$f(\tau)=\bar{j}^2x^2(1-P)[(1-\operatorname{erf}y)+2y\pi^{-\frac{1}{2}}\exp(-y^2)].$$

The spectral power density of the noise current (using Khintchine's theorem) is then

$$R(\omega)=4\int_0^{\infty} f(\tau)\cos\omega\tau\cdot d\tau$$
$$=4\bar{j}^2x^2[(1-\operatorname{erf}y)+2y\pi^{-\frac{1}{2}}\exp(-y^2)]F(P),$$

where

$$F(P)=\int_0^{\infty}[1-P(\tau)]d\tau.$$

Macfarlane evaluates $F(P)$ for the particular cases of circular and long thin patches, so obtaining the frequency characteristic of the mean-square noise current $\langle\delta j\rangle_{Av}^2$.

We are here concerned with the dependence of $\langle\delta j\rangle_{Av}^2$ on the mean emission current $\bar{j}$ and on the magnitude of the retarding

field. For a given j_0 and as y varies, one obtains

$$\langle \delta j \rangle_{Av}^2 \propto j^2[(1-\text{erf}y)+2y\pi^{-\frac{1}{2}}\exp-y^2].$$

For $y>0.4$ the mean-square noise current falls off more rapidly than j^3 and becomes very small for $y>2$. This implies that in the retarding field the mean-square noise current falls off more rapidly than is given by the square law relation, and this is contrary to practical experience.

One concludes that this model for flicker noise is not suitable in its present form although qualitatively it is a more likely explanation of the effect (considered as a surface phenomenon) than any of the other theories mentioned above. The conception of local increases of the potential barrier at the cathode surface is not contradictory to the generally accepted theory of the presence of free barium on the crystal surfaces. Local poisoning effects, for instance due to the recombination of positive barium with negative ions, would cause a large increase of work function in agreement with the physical picture outlined above. As far as changes of effective work function are concerned, there is no essential difference between the migration of negative ions on to a patch or, conversely, the movement of positive barium ions out of the patch. Equally, the model for flicker noise may be assumed to account for the migration of positive barium ions over a region of high-value work function, so causing an increase of emission from that patch. To be consistent with the theory, the emission must simply be restricted to values lying between zero and that value given by the initial velocity distribution and the negative anode-cathode potential difference; this value cannot be exceeded whatever the reduction of effective work function at the cathode surface.

[1] W. Schottky, Phys. Rev. **28**, 74 (1926).
[2] G. G. Macfarlane, Proc. Phys. Soc. (London) **59**, 366 (1947).
[3] A. Van der Ziel, Physica **16**, 359 (1950).
[4] F. K. du Pre, Phys. Rev. **78**, 615 (1950).
[5] T. B. Tomlinson, J. Appl. Phys. **23**, 894 (1952).
[6] G. G. Macfarlane, Proc. Phys. Soc. (London) **63**, 807 (1950).

Noise in CW Magnetrons*

DAVID MIDDLETON, *Harvard University, Cambridge, Massachusetts*
W. M. GOTTSCHALK, *Research Division, Raytheon Manufacturing Company, Waltham, Massachusetts*
AND
J. B. WIESNER, *Massachusetts Institute of Technology, Cambridge, Massachusetts*
(Received May 11, 1953)

A STUDY of CW magnetrons has been undertaken in an attempt to measure and control the noise accompanying the carrier oscillations and to provide a more complete understanding of the ultimate physical mechanisms involved. The first problem has been to determine the correct macroscopic model of the tube's output, e.g., carrier plus noise, carrier amplitude-modulated by noise, or carrier angle-modulated by noise, etc.—from measurements made with suitably designed apparatus. The choice of the proper model is essential for a correct interpretation of the observed data, since all measuring equipment of necessity contains at least one effectively nonlinear element—*viz.*, a rectifier or detector. Microwave AM and FM receivers, and direct detection systems (of the type described below) have been built for this purpose, designed to fit closely the somewhat idealized conditions required for the analysis of signals and noise in nonlinear systems.[1,2] The preliminary results obtained by the direct detection scheme not only have indicated the proper macroscopic model of the magnetron's cw output, but have also given us quantitative measures of the model's structure as well.

The CW magnetrons tested have been chiefly the type RK5609, whose normal power output is about 100 watts. This tube has an oxide-coated cathode and a coaxial type output. For the experimental work the heater has been battery operated and a well-filtered, current-regulated power supply has been used. The magnetrons are operated into a matched load (with a voltage SWR less than 1.1), and data are obtained by the direct detection method shown in the diagram (Fig. 1). In this form the tunable cavity acts like a "slope" detector, or discriminator with respect to angle-modulated inputs.

When the magnetron and other apparatus have reached an equilibrium state, a typical procedure requires tuning first the cavity to the magnetron frequency f_c; the probe and video attenuator are then adjusted for suitable indications of the milliammeter (I_{dc}) and thermocouple (P_c). The cavity (with central frequency f_0) is then detuned in one direction by an amount $f_D=f_c-f_0$, and f_D is accordingly varied point by point over several widths ($\omega_F/2\pi$) of the wave-guide filter. At each setting the steady component of crystal current and the thermocouple reading are recorded.

Two quantities measured so far by this method are $P_{dc}(\sim I_{dc}^2)$ and P_c, the mean power of the fluctuating continuum observed in the video. A variety of possible macroscopic models of the magnetron's output have been tested and the results indicate† that for these tubes the predominant effect is that of a carrier (at f_c), angle-modulated by a band of random noise. (There is also a relatively small amount of amplitude-modulation by noise, as well

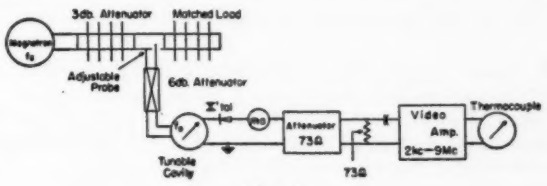


FIG. 1.

as an accompanying noise background or "shot" effect, both of which we expect from physical considerations.) Analysis based on the above model in the usual strong-carrier conditions shows[3]

$$P_c/P_{dc}=4\pi^2K^2(\omega_D)\langle(d\Phi/dt)^2\rangle,$$

where $K(\omega_D)=[|Y_0|^{-1}d|Y_0|/d\omega_D]$, Y_0 is the slowly varying part of the system function for the cavity, and $d\Phi/dt$ is the frequency deviation of the carrier. Typical curves are illustrated in Fig. 2; as expected, the rms deviation varies considerably with operating conditions. With these tubes, for example, it was found to be 12.7 kc/sec±1.5 kc/sec at a plate-current level $I_b=125$ ma. At $I_b=150$ ma, $\langle(d\Phi/dt)^2\rangle^{\frac{1}{2}}=10.7$ kc/sec±1.1 kc/sec for a typical run. The order of magnitude in all cases so far has been 10–20 kc/sec.

Frequency deviations of the magnitude measured in these experiments can be accounted for by taking into account the frequency-modulation caused by fluctuations in the density of the space charge present in the magnetron's interaction space. These fluctuations may be due to flicker noise or other cathode noise phenomena, or possibly to the presence of gas ions in the tube.

The effects of the electron cloud in the interaction space can be characterized by an equivalent shunt electronic admittance in parallel with the tuned circuit of the magnetron.[4] The reactive components of the electronic admittance, the electronic susceptance b_{el}, will alter the natural frequency of the resonant structure of the tube. b_{el} itself is a function of the anode current, magnetic field, rf voltage, and external load impedance, as well as the tube design; thus the actual frequency will be a complicated function of these parameters.[5,6] Fluctuations in b_{el} produce corresponding fluctuations in the instantaneous frequency of the tube. The observed rms frequency deviation, $\langle(d\Phi/dt)^2\rangle^{\frac{1}{2}}$, can be accounted for by the fluctuations in b_{el}. In the simple theory b_{el} is assumed to be proportional to the number of electrons present in the interaction space, and the rms fluctuation in the number of electrons is assumed to be equal to $\sqrt{n}$, where n is the actual number of electrons present. Calculations based on these assumptions yield values for the rms frequency deviation $\langle(d\Phi/dt)^2\rangle^{\frac{1}{2}}$ ranging from 10 cps to 10 000 cps. The actual value of $\langle(d\Phi/dt)^2\rangle^{\frac{1}{2}}$ depends upon the rate of change of frequency with anode current

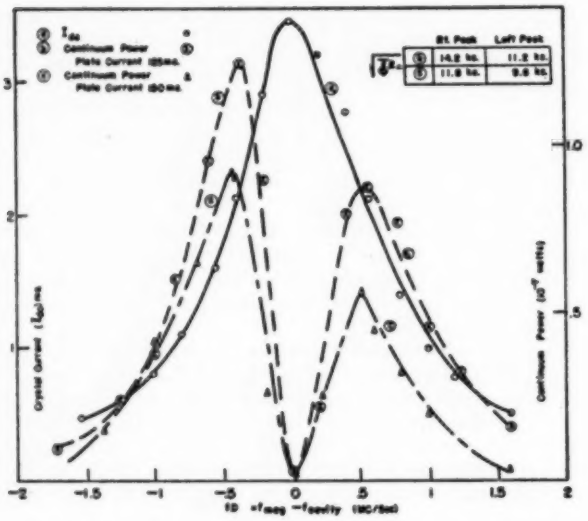


FIG. 2. Direct detection data for magnetron RK5609 No. C4818D.

(the instantaneous pushing figure df/dI_a, at the operating point) and the total number of electrons in the interaction space.

The FM noise power is expected to have its least value at those points on the operating characteristic when df/dI_a is a minimum. At the present time an attempt is being made to correlate df/dI_A and the FM noise output for the tube discussed above.

The fluctuations in current density should also show an effect in the dc supply and a high correlation is expected between frequency variations and fluctuations in the anode current. Correlation coefficients ranging from approximately 0.80 to 0.95 have been measured when the video output of the FM detector was compared with the video signal produced across a small resistance in series with the magnetron.

Narrow-band frequency modulation of a carrier by a wide-band noise produces a modulated signal having no carrier and a very narrow continuous spectrum. For the tube under investigation the spectral width at the half-power points is expected to have values in the range between ½ cps and 50 cps, depending upon the operating point of the tube and the exact character of the modulating noise.[7]

* The work reported in this note is part of a study carried out at the Research Division of Raytheon, under Air Force Contracts AF19(122)-473 and AF19(604)-636, sponsored by the Air Force Cambridge Research Center, Cambridge, Massachusetts.

† Details are given in Progress Report No. 1, May 1, 1953.

[1] S. O. Rice, Bell System Tech. J. **23**, 282 (1944); **24**, 46 (1945).
[2] D. Middleton, Quart. Appl. Math. **5**, 445 (1948); for angle-modulation, Phil. Mag. Ser. 7, **42**, 689 (1951); (and demodulation) J. Appl. Phys. **20**, 334 (1949) Quart. Appl. Math. **7**, 129 (1949); **8**, 59 (1950).
[3] Quarterly Reports Nos. 5, 6, *Microwave Noise Study*, Raytheon Manufacturing Company, section 6, (March 17–September 17, 1952).
[4] J. C. Slater, *Microwave Electronics* (D. Van Nostrand Company, Inc., New York, 1950).
[5] G. B. Collins, *Microwave Magnetrons*, McGraw-Hill Book Company, Inc., New York, 1948).
[6] H. W. Welch, Jr., Technical Report No. 1, Electron Tube Laboratory, Department of Electrical Engineering, University of Michigan.
[7] Quarterly Report No. 6, *Microwave Noise Study*, section 7, Research Division, Raytheon Manufacturing Company, September 17, 1952.

Effect of Halogenation on Frictional Properties of Plastics

R. C. BOWERS, W. C. CLINTON, AND W. A. ZISMAN
Naval Research Laboratory, Washington, D. C.
(Received April 22, 1953)

THIS is to report recent results on the friction and wear properties of unplasticized high polymers. Plastics studied were polymers of ethylene and halogenated substituents, including: polyethylene, polyvinyl chloride, polyvinylidene chloride, polyvinyl fluoride, a 1:1 copolymer of ethylene with tetrafluoroethylene, polytetrafluoroethylene (TFE), polytrifluorochloroethylene (TFCE), and a series of five copolymers of the last two. Coefficients of static friction (μ_s) and of kinetic friction (μ_k) were measured at 25°C on a "stick-slip" machine[1] with a sliding speed of 0.01 cm/sec, loads from 0.2 to 1.5 kg, and steel balls from 6.35 to 50.7 mm in diameter. Observations were made for steel sliding on plastic, plastic on steel, and plastic on plastic. Our values of μ_s for TFE, polyvinyl chloride, and polyethylene agreed with those of Shooter and Tabor.[2] Except for TFE, our measured values of μ_s for steel sliding on all plastics were approximately equal to the ratio of the shear strength to the yield pressure of the plastic. The exception arose from the existence of considerable slipping at the interface of metal and TFE.

Smooth, heat-cast plastics exhibited erratic and significantly higher values of μ_s and μ_k than did the abraded materials. This was caused by the presence of a thin, softer, and more amorphous surface layer. Rotary abrasion under water removed it, exposing a surface characterized by reproducible values of μ_s and μ_k. Transfer of plastic to steel was the principal cause of the change in μ_k during sliding and of the difference between μ_s and μ_k. Polyethylene and TFE transferred most readily to form adherent films which electron diffraction proved were highly oriented in the direction of sliding. Friction, wear, and cold-welding were greater the higher the proportion of substituted chlorine, and they were smaller the higher the proportion of fluorine. Thus, values of μ_s of 0.10, 0.30, 0.33, 0.45, and 0.69 were found for steel sliding on abraded TFE, 1:1 copolymer of tetra-fluoroethylene and polyethylene, polyvinyl chloride, and polyvinylidene chloride, respectively.

The increased substitution of fluorine in polyethylene lowered μ_s and μ_k only slightly for the first 50 atom percent substitution but rapidly decreased them as 100 percent substitution was approached. The substitution of chlorine for fluorine in TFE had a remarkably large effect on friction. A rise of 4-fold in μ_s and 6-fold in μ_k occurred from the replacement of only 10 percent of the fluorine atoms.

Our conclusions parallel those made recently[3–6] concerning the effect of halogen substitution on wettability, liquid adhesion, and the free surface energy of polymeric solids. The van der Waals forces between the polymer molecules, as well as those between polymer molecules and molecules of contacting liquids and solids, appear to obey the same qualitative rules relative to the effect of chlorine or fluorine substitution. For this reason the internal cohesional, adhesional, and frictional properties increase with chlorine substitution and decrease with fluorine substitution. A similar change in these properties can be predicted when covalent hydrogen is replaced by fluorine or chlorine in any polymer.

Friction measurements were made on polished steel plates coated with thin films of these polymers, and μ_s and μ_k were always considerably less than the values for the corresponding bulk plastics. Thus, with polyethylene rubbed on steel, μ_s was 0.05, and in bulk it was 0.33. The cause was the decrease in the real area of contact resulting from the presence of the harder steel substrate. The mean yield pressure of the steel substrate and of each plastic was measured, and it was found that the product of the coefficient of friction of the bulk plastic and the ratio of the yield pressure of the bulk plastic to that of the steel substrate was approximately equal to the observed coefficient of friction of the coated steel.

Durable and adherent films of TFE on steel and brass were prepared either by oven baking or by high-frequency induction heating with an aqueous suspensoid of TFE. Films from 0.0005 in. to 0.005 in. thick were studied. Values of μ_s and μ_k for steel sliding on such coatings were always 0.04, and this was caused by immediate transfer of polymer to give the friction of TFE against TFE. Ease of material transfer, combined with low yield pressure and low shear strength, was confined to polyethylene and TFE. Hence, these polymers, alone of those studied, invite applications as solid lubricants. The additional properties of extremely low adhesion and high heat stability make TFE particularly valuable

as a solid lubricant. The durability of such films of TFE was shown by making 100 successive traverses of a loaded steel ball on a coated steel plate with no visual evidence of wear and only a slow rise of μ_k from 0.04 to a final value of 0.08.

[1] F. P. Bowden and D. Tabor, *The Friction and Lubrication of Solids*, (Oxford University Press, London, England, 1950).
[2] K. V. Shooter and D. Tabor, Proc. Phys. Soc. (London) **B65**, 661 (1952).
[3] H. W. Fox and W. A. Zisman, J. Colloid Sci. **5**, 514 (1950).
[4] H. W. Fox and W. A. Zisman, J. Colliod Sci. **7**, 109, 428 (1952).
[5] F. Schulman and W. A. Zisman, J. Colloid. Sci. **7**, 465 (1952).

Comment on the Temperature Dependence of the Viscosity of Liquids

ARTHUR K. DOOLITTLE
Carbide and Carbon Chemicals Company, South Charleston, West Virginia
(Received March 13, 1953)

IN an article[1] published in this Journal in September, 1952, Gutmann and Simmons disagree with statements made by the writer in an earlier paper[2] that "even the best of the relatively simple viscosity–temperature functions is inadequate to reproduce the measured values accurately over extended ranges of temperature," and "relationships that reproduce the measured values with satisfactory validity are necessarily complicated and not likely to hold promise of fundamental significance."

In support of their contention they cite the Vogel[3] equation (without credit either to Vogel who originated this equation in 1921, or to Herschel,[4] Cragoe,[5] or Hugel,[6] all of whom worked with it extensively).

Gutmann and Simmons employ the Gaussian criterion of fit to prove that the Vogel equation

$$\ln\eta = A + B/(T+C) \qquad (1)$$

gives a better fit of our *n*-paraffin data than does our more complicated equation

$$\ln\eta = -4.54e^{500/T}[-Ei(-500/T)] - 4.66 + C_1 e^{500/T}. \qquad (2)$$

We refute this argument by pointing out that the Gaussian criterion of fit is not applicable to dissimilar equations. This criterion was developed for the purpose of determining the number of terms that are justified in a power series and is not appropriate to a comparison of Eqs. (1) and (2).

While we did not claim (in our article) that the above equation [Eq. (2)] reproduced the measured values with *satisfactory* validity, we did say that the calculated values of the viscosity (of the *n*-paraffins) were as close to the observed values as the accuracy of the measurements in 65 percent of the cases reported. Gutmann and Simmons, on the other hand, say "Eq. (1) is quite adequate in representing temperature–viscosity data over substantial ranges in temperature," and "for all the pure liquids listed (in their Table II) the Standard Percentage Errors (SPE) *lie within the experimental errors*." This statement by them is false in so far as it applies to our data on the *n*-paraffins, since in no case where their equation is applied over a substantial range of temperatures (at least 200 centigrade degrees) is the SPE as low as the accuracy of our measurements.

In spite of our reluctance to claim that Eq. (2) reproduces the measured values with *satisfactory* validity, it is nevertheless readily shown that Eq. (2) gives a better fit than the Vogel equation [Eq. (1)] when applied to data over an extended range of temperatures. Measurements on only three of our *n*-paraffins were made over ranges in excess of 200 centigrade degrees, hence these three compounds only (undecane, tridecane, and heptadecane) will be considered in making the comparison. Since the lowest temperature values reported by us for both tridecane and heptadecane were extrapolated below the freezing points of these compounds, we show both the extreme range and a range in which the lowest temperature point is rejected, in Table I. The criterion of closeness of fit is the Standard Percentage Error.

TABLE I. Comparison of Eq. (1) and Eq. (2) for undecane, tridecane, and heptadecane.

Compound	Temperature Range	*n*	Accuracy	SPE Eq. (1)	SPE Eq. (2)
n-undecane	263°–473°K	6	0.36	1.23[a]	0.62
n-tridecane	263°–473°	6	0.35(5 pts)	1.38	2.34
n-heptadecane	293°–573°	7	0.43	2.24[a]	1.72
n-tridecane	293°–473°	5	0.35	0.60	0.55
n-heptadecane	323°–573°	6	0.43	1.39	0.65

[a] These values were erroneously calculated by Gutmann and Simmons in their Table II to be 1.5 and 1.8, respectively. It should also be mentioned that the values listed in their Table II ascribed to *n*-heptane are those of *n*-octane.

It is clear from the above tabulation that in all cases where the data are appropriate, Eq. (2) gives a better fit than Eq. (1). Only when a value of tridecane at −10°C is employed does Eq. (1) appear to be the better. This value is extrapolated from a measurement made at −5.48°C. Attention is called to the fact that neither equation represents the data as accurately as the measurements.

In reference to the attempt by Gutmann and Simmons to derive the Vogel equation [Eq. (1)] from the Eyring equation, we need merely comment that it would be just as valid to assume Eq. (1) itself as to assume that $E = E_0 + f(T)$.

[1] F. Gutmann and L. M. Simmons, J. Appl. Phys. **23**, 977 (1952).
[2] A. K. Doolittle, J. Appl. Phys. **22**, 1031 (1951).
[3] H. Vogel, Physik Z. **22**, 645 (1921).
[4] W. H. Herschel, Oil Gas J. **25**, 146 (1926).
[5] C. S. Cragoe, Proc. World Petroleum Cong. (London) **2**, 529 (1933).
[6] G. Hugel, Oel u. Kohle ver. Erdoel u. Teer **15**, 27 (1939).

Further Note on the Temperature Dependence of the Viscosity of Liquids

F. GUTMANN AND L. M. SIMMONS
The New South Wales University of Technology, Broadway, Sydney, Australia
(Received May 25, 1953)

AT no time have we claimed originality for the viscosity relationship

$$\log\eta = A + B/(T+C), \qquad (1)$$

but statements in Doolittle's paper[1] made it desirable to call attention to the existence of this equation. Although the literature cited in our publication[2] was, of course, not a bibliography, it did refer to three authors who had previously used this equation.

The important thing is that the equation exists, is simple, has a closed form, does not require recourse to mathematical tables beyond ordinary logarithms, and adequately expresses over extended temperature ranges the viscosity data for all those liquids quoted by us[2] with the possible exception of three of those liquids for which Doolittle claims the remarkably high accuracy of less than 0.5 percent. It may be of interest to note that subsequent to our previous publication,[2] we have examined Thorpe and Rodger's[3] viscosity-temperature data for formic and acetic acids, and methyl, propyl, and butyl alcohols, and find that for them the SPE's are less than 0.5 percent. As far as we know, greater accuracy has not been claimed for these data, the temperature ranges of which are respectively 281 to 370, 304 to 386, 277 to 336, 280 to 369, and 273 to 387°K.

With reference to the applicability of the Gaussian criterion of fit (Ω), we quote Worthing and Geffner:[4] "Of all equations, that which yields the least value of Ω is assumed to best represent the given data." Another statistical means for finding whether the improvement in fit on introducing an additional free parameter is due to other than mere chance is Fisher's "*F*-Test."[5] Applying this test, we find that only in the case of *n* heptane is the probability less than 5 percent that the improvement in fit, obtained by the use of Doolittle's 4-parameter equation instead of the 3-parameter equation (1), is due to chance.

We cannot accept the statement that "it will be just as valid to assume Eq. (1) itself as to assume that $E=E_0+f(T)$." This is the most general statement of E as an analytic function of T, but Eq. (1) is certainly not the most general statement of η as an analytic function of T. Furthermore, $f(T)$ is known to be a small correction term in a relatively large activation energy, and so it is logical, as well as most general, to express the activation energy as the sum of a constant and a temperature dependent term.

[1] A. K. Doolittle, J. Appl. Phys. **22**, 1031 (1951).
[2] F. Gutmann and L. M. Simmons, J. Appl. Phys. **23**, 977 (1952).
[3] T. E. Thorpe and J. W. Rodger, Trans. Roy. Soc. (London) **185**, 397 (1894).
[4] A. G. Worthing and J. Geffner, *Treatment of Experimental Data* (John Wiley and Sons, Inc., New York, 1943) p. 266.
[5] R. A. Fisher and F. Yates, *Statistical Tables* (Oliver and Boyd, London, 1948).

Summation of Slowly Converging Series*

IGOR GUMOWSKI
Université Laval, Québec, Canada
(Received May 4, 1953)

SLOWLY converging series occur quite frequently in problems of pure and applied science. In general, these series are not suitable for numerical computations. However, following up an idea suggested by A. Robert,† we found that the following method permits a rapid summation of such series.

Consider a slowly converging series,

$$S=\sum_{k=1}^{\infty} a_k. \tag{1}$$

If

$$f(n)=\sum_{k=1}^{n} a_k,$$

then

$$f(n+1)-f(n)=a_{n+1}. \tag{2}$$

Letting

$$a_{n+1}=p'(n+1),$$

we can write (2) as follows:

$$f(n+1)-f(n)=p'(n+1). \tag{3}$$

If we can solve the difference equation (3) for $f(n)$, the sum of the series (1) will be given by

$$S=\lim_{n\to\infty} f(n).$$

Since $f(n+\lambda)=e^{\lambda D}f(n)$, where $D=d/dn$ and λ is a constant, the difference equation (3) can be written

$$(e^D-1)f(n)=De^D p(n).$$

Taking Laplace transforms of both sides and solving for $f(n)$, we obtain

$$f(n)=\frac{1}{2\pi j}\int_{Br}\frac{e^{nz}e^z z g(z)}{e^z-1}dz, \tag{4}$$

where Br is a Bromwich contour and

$$g(z)=\int_{0^-}^{\infty} e^{-zn}p(n)dn.$$

Because

$$\int_{0^-}^{\infty} e^{-zn}D^m p(n)dn=z^m g(z)$$

and we can write

$$\frac{ze^z}{e^z-1}=\sum_{m=0}^{\infty} b_m z^m, \tag{5}$$

Eq. (4) becomes

$$f(n)=\frac{1}{2\pi j}\sum_{m=0}^{\infty} b_m\int_{Br} e^{zn}z^m g(z)dz$$

$$=\sum_{m=0}^{\infty} b_m D^m p(n),$$

which is the "particular integral" of the difference equation (3). The complete solution of the difference equation is given by

$$f(n)=A+\sum_{m=0}^{\infty} b_m D^m p(n), \tag{6}$$

where A is a constant or a periodic function of period 1. Since the series (1) is convergent, $\lim_{n\to\infty} a_n=0$ and $\lim_{n\to\infty} D^m a_n=0$, for $m=1, 2, \cdots$; therefore

$$S=A+\lim_{n\to\infty}\sum_{m=0}^{\infty} b_m D^m p(n)$$

$$=A+b_0\lim_{n\to\infty} p(n).$$

But

$$p(n)=\int_c^n a_n dn,$$

where c is a constant, therefore

$$S=A+b_0\int_c^{\infty} a_n dn. \tag{7}$$

Solving (6) for A and replacing in (7), we obtain

$$S=f(n)+b_0\int_c^{\infty} a_n dn-b_0 p(n)-\sum_{m=1}^{\infty} b_m D^m p(n)$$

$$=f(n)+b_0\int_n^{\infty} a_n dn-\sum_{m=1}^{\infty} b_m D^m p(n). \tag{8}$$

Examining (8) we notice that

$$S=f(n)+\text{remainder},$$

where the remainder is a rapidly converging series.

Since the constants b_m are independent of a_n, they need be calculated only once. From (5) we can write immediately

$$b_m=\frac{1}{m!}\lim_{s\to 0} D^m\left(\frac{ze^z}{e^z-1}\right).$$

The first seven constants have the following values:

$$b_0=1,\quad b_1=\tfrac{1}{2},\quad b_2=\frac{1}{2!}\cdot\frac{1}{12},\quad b_3=0,\quad b_4=-\frac{1}{4!}\cdot\frac{1}{720},$$

$$b_5=0,\quad b_6=\frac{1}{6!}\cdot\frac{1}{30\,240}.$$

Examples:

(a) $$S=\sum_{k=1}^{\infty}\frac{1}{k^2}.$$

In order to sum this series with a precision of five decimals approximately 200 000 terms are necessary. Letting $n=10$ and using (8) we obtain

$$p(n)=1/n=0.1000000,\quad f(n)=1.5497678,\quad b_1p'(n)=0.0050000,$$
$$b_2p''(n)=0.0001667,\quad b_4p''''(n)=0.0000003.$$

Therefore, $S=1.644934$. However the sum of this particular series is known, it is equal to $\pi^2/6$, which according to the tables is equal to 1.644934.

(b) $$S=\frac{1}{\pi}\sum_{k=1}^{\infty}\frac{\sin kx}{k}.$$

The relation (8) for this series yields $(0\leq x\leq\pi)$

$$S=\frac{1}{\pi}\sum_{k=1}^{n}\frac{\sin kx}{k}+\frac{1}{\pi}\int_{nx}^{\infty}\frac{\sin y}{y}dy-\frac{\sin nx}{2n\pi}-\frac{x\cos nx}{24n\pi}-\cdots.$$

We notice that the above expression is quite easy to sum. Choosing again $n=10$ and calculating with three decimals only, we found that S represents the function

$$f(x)=\tfrac{1}{2}-(x/2\pi),\quad (0\leq x\leq\pi).$$

* This contribution is a part of a D.Sc. thesis presented at the Department of Electrical Engineering of Laval University in Quebec, Canada.
† Presented at an informal gathering at Laval University.